②

AD-A229 621

DTIC
ELECTE
DEC 21 1990
S E D

DEPARTMENT OF THE AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

90 12 20 106

AFIT/GSM/LSY/90S-18

MISLEADING GRAPHICS: CAN
DECISION MAKERS BE
AFFECTED BY THEIR USE?

THESIS

Albert A. Larkin, Capt, USAF

AFIT/GSM/LSY/90S-18

Approved for public release; distribution unlimited

The opinions and conclusions in this paper are those of the author and are not intended to represent the official position of the DOD, USAF, or any other government agency.

Accession For		
NTIS GRA&I		X
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A-1		

AFIT/GSM/LSY/90S-18

MISLEADING GRAPHICS: CAN DECISION MAKERS

BE AFFECTED BY THEIR USE?

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Systems Managemen.

Albert A. Larkin, B. S.

Captain, USAF

September   1990

Table of Contents

List of Figures

List of Tables

AFIT/GSM/LSY/90S-18

Abstract

This thesis investigated how Air Force decision makers can be affected by misleading graphics. A literature review revealed the existence of criteria for creating high-integrity graphs, as well as style guides for formatting and procedural techniques. By violating the criteria, misleading graphs may be produced which misrepresent the underlying data. Experiments were conducted on 63 Air Force Institute of Technology short-course students. Graphs typically used in the analysis of cost performance reports were constructed in a way to violate one of the criteria for the experiments. Using a t-test, it was demonstrated that Air Force decision makers can be misled by graphs that violate the criteria at a level of significance of less than .0001. Also, a sampling of graphs from throughout Air Force Systems Command revealed that program managers and cost analysts were creating graphs that violated the criteria.

MISLEADING GRAPHICS: CAN DECISION MAKERS
BE AFFECTED BY THEIR USE?

I. Introduction

General Issue

Many times in everyday situations, data are presented in a

tabular format. This format does not readily lend itself to easy

data comprehension, especially when several variables are

presented. While some variables may be increasing over a period

of time, some may be decreasing, and others may be constant.

People using data like this often need a quick way to analyze it.

Graphing often provides a relatively fast and easy means of

analysis. However, graphs may easily be made to portray the data

in a more (or less) favorable light.

For example, the recent volatility in the Dow Jones

Industrial Average (DJIA) could be graphed over some time period,

perhaps one year, showing in great detail the wide fluctuations.

By extending the vertical scale beyond the highest data point,

these recent fluctuations may be "smoothed out" and their

relative impact could be minimized. To illustrate, Figure 1

represents the DJIA for the year of 1989, while Figure 2

represents the same data but plotted against a much larger scale.

Clearly, the expanded scale minimizes the impact of the

volatility.

DJIA 1989 CLOSING FIGURES
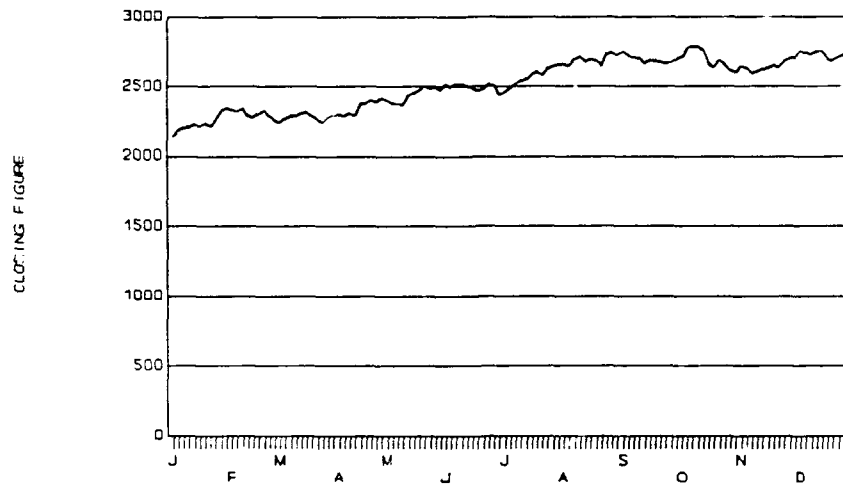
Figure 1. High-Integrity Graph

DJIA 1989 CLOSING F'GURES
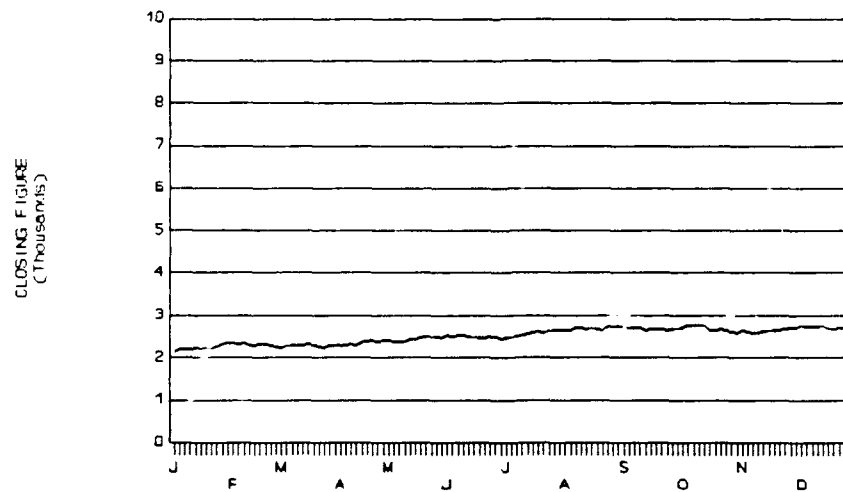
Figure 2. Extended Vertical Axis

2

Changing a scale is only one method to make graphs misleading; many other methods exist and are used everyday in newspapers, magazines, technical literature, etc. (Appendix A contains several examples of graphs found in common reading materials that distort the effect the underlying data are suggesting).

Academic literature identifies criteria for constructing high integrity graphs, and asserts that the use of misleading or distorted graphs may lead to faulty decisions. Returning to the DJIA example, an unscrupulous stockbroker could minimize a prospective client's concern over the stock market's volatility by showing the client the extended scale DJIA graph. The client may then decide the stock market is relatively stable, when in fact the market may have been judged to be very erratic had the client viewed Figure 1. The results of the client's decision to enter the stock market could be very damaging to him, because he was reliant on misleading information.

Specific Problem

Cost Performance Reports (CPRs) are a key tool used by System Program Offices (SPOs) to determine a contractor's cost and schedule status on a project. CPRs are required submittals on major weapon systems contracts, and contain much data, in tabular format, detailing the amount of progress made toward the budgetary and schedule goals of the project. An example of a CPR is shown in Figure 3. The

COST PERFORMANCE REPORT - WORK BREAKDOWN STRUCTURE

CONTRACT | PROGRAM '86/87 BUY PROD | REPORT PERIOD 09-30-88 TO 10-27-88

DOLLARS-IN-MILLIONS

WORK BREAKDOWN STRUCTURE ITEM	CURRENT PERIOD						CUMULATIVE TO DATE						AT COMPLETION		
	BUDGETED COST		ACTUAL COST WORK PERFORMED	VARIANCE			BUDGETED COST		ACTUAL COST WORK PERFORMED	VARIANCE			BUDGETED	LATEST REVISED ESTIMATE	VARIANCE
	WORK SCHEDULED	WORK PERFORMED		SCHEDULE	COST		WORK SCHEDULED	WORK PERFORMED		SCHEDULE	COST				
SUBTOTAL															
AGING & SURVEILLANCE															
DIRECT LABOR															
DIRECT MATERIAL															
OTHER DIRECT COSTS															
SUBTOTAL															
LIFE CYCLE COST 3.10.10															
DIRECT LABOR															
OTHER DIRECT COSTS															
SUBTOTAL															
ECP-5															
DIRECT LABOR															
DIRECT MATERIAL															
OTHER DIRECT COSTS															
SUBTOTAL															
SYS ENG MGMT-SYS ENGRG															
DIRECT LABOR															
DIRECT MATERIAL															
OTHER DIRECT COSTS															
IOWA															
SUBTOTAL															
TECHNICAL PUBLICATIONS															
DIRECT LABOR															
OTHER DIRECT COSTS															
SUBTOTAL															
DATA															
DIRECT LABOR															
DIRECT MATERIAL															

Figure 3. Cost Performance Report

4

tabular format can make the data difficult and time consuming to comprehend. SPO program managers and analysts may plot the data to provide a graphical representation. By plotting the CPR data, a "picture" is drawn which can reveal much about the project's cost and schedule trends that the tabular format does not readily divulge.

The availability of desktop computers and graphing software (i.e, Quattro, Lotus 1-2-3, Harvard Graphics, Enable, etc.) throughout Air Force Systems Command (AFSC) enables SPO managers and analysts to easily plot CPR data. While this can be helpful to the manager/analyst, they may be (consciously or unconsciously) plotting the data in such a manner as to be misleading and potentially damaging. An example of a misleading graph is displayed at top in Figure 4, while the high-integrity version of the same data is shown at the bottom. Upon close observation, one can see the more irregular stratum (direct costs) has been placed at the bottom. This creates the illusion that the indirect costs are decreasing, while they have been constant throughout the period. Faulty managerial decisions may result from using graphs like these.

Consider a program in which the contractor is over budget (ACWP > BCWP). BCWP and ACWP are defined in Chapter II. In an attempt to mask this unfavorable condition, an unscrupulous analyst could create a graph with two vertical axes, such that the BCWP appears greater than ACWP (at top of Figure 5). An unsuspecting decision maker may be misled
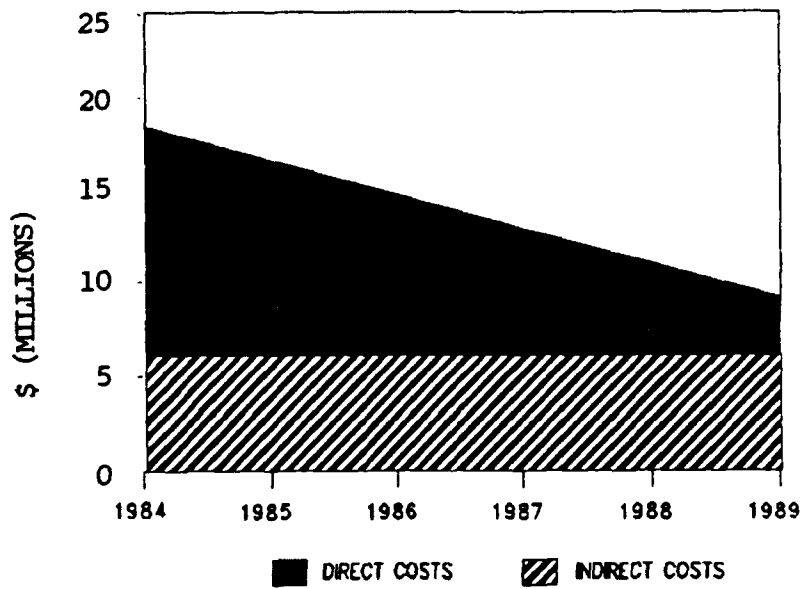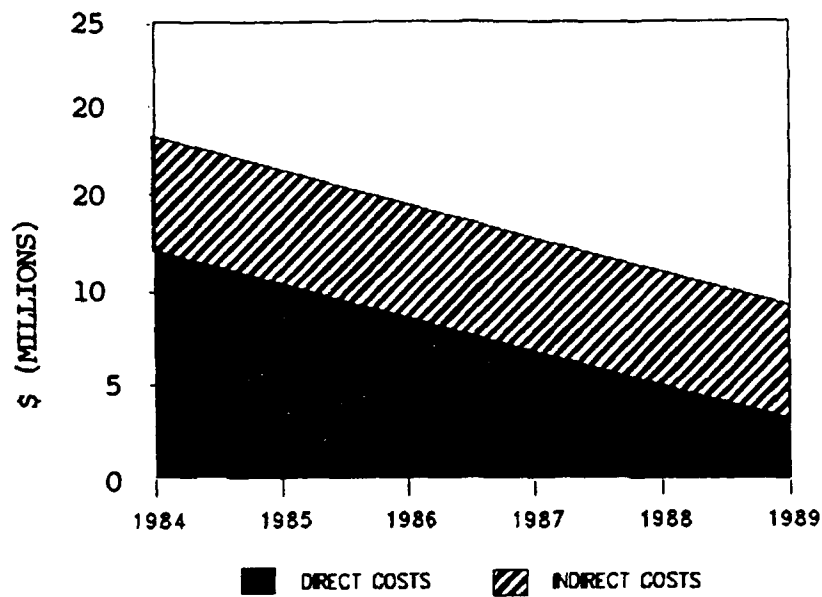
Figure 4. Misleading (top) and High-Integrity
(bottom) Graphs of the Same Data.
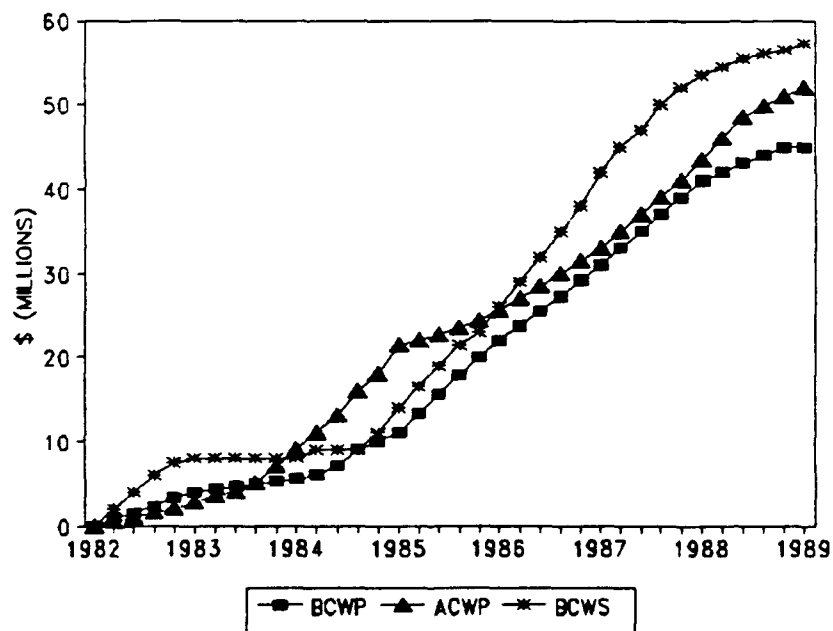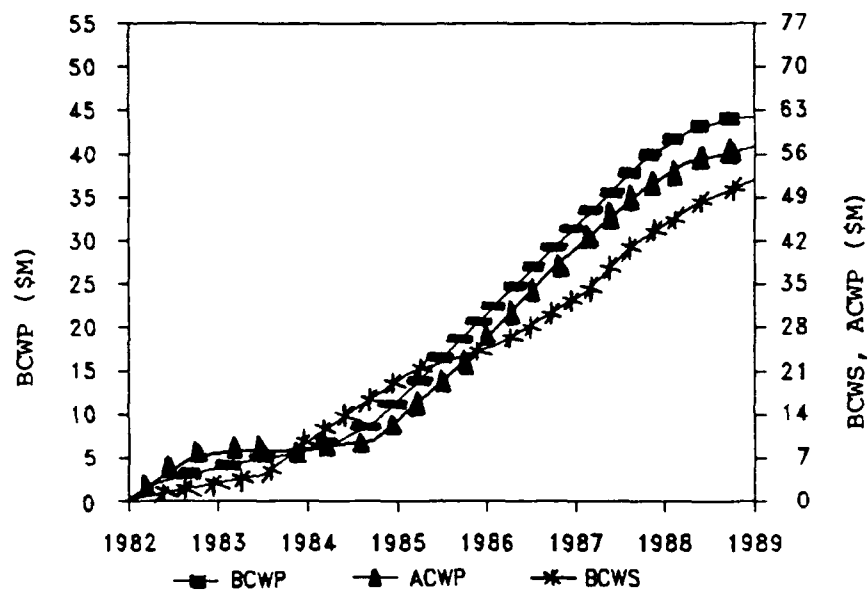
Figure 5. High-Integrity and Misleading Versions of a
Cumulative Contract Status Graph.

by such a graph. The high-integrity version of the same data appears at the bottom of Figure 5. In some instances, an analyst may be completely unaware of the distortion created by manipulating graphics software. However, the danger of faulty decision making based on these graphs is still just as ominous.

The Air Force cannot afford to make faulty decisions, especially now that defense spending is being significantly reduced. Every dollar must be stretched as far as possible. To support proper decisions with limited financial resources, the information presented to our decision makers must be completely truthful and free of distortion. In our "fishbowl" environment, in which defense spending is closely monitored by the Congress, the public, various lobbies, and other interested parties, costly mistakes simply will not be tolerated. The correct decisions must be made the first time.

Objectives

The objectives of this thesis are two-fold: determine what graphs of CPR data are used to formulate managerial decisions, and determine if distorted graphs of CPR data can lead to faulty decisions. The investigative questions to be used are:

1. What kinds of graphs of CPR data are being used?

2. Are there criteria, principles, or guidelines of graphical excellence?

3. Are program managers/analysts creating graphs which violate the criteria?

4. Can graphs of CPR data which violate the criteria mislead decision makers?

Limitations

There are limitations to this thesis that warrant discussion. First, the experiment was conducted in a rigidly controlled classroom setting. Although this limits external validity, it strengthens internal validity. Second, the time allowed to review graphs and respond to conclusions based on those graphs was fixed to simulate the time a busy program manager or cost analyst would have to review graphs and make decisions. Mintzberg has characterized the managerial world as one of brevity, in which managers rarely spend more than 30 minutes on any one activity (1:55). Allowing brief time periods to review graphs is therefore not unreasonable.

Conclusion

The methods used to answer the investigative questions are discussed in Chapter II, Literature Review, and Chapter III, Methodology. Chapter II contains tables of criteria for producing high-integrity graphs, along with style guides to help make the graphs more clear and legible. In addition, one method for determining the amount of distortion contained in a graph is discussed. Chapter III discusses the construction and execution of an experiment to determine if decision makers could be misled. Chapter IV,

Analysis and Findings, contains all findings involved in the research of these objectives. Chapter V contains a summary of all findings, along with recommendations for further research.

II. Literature Review

The literature review is comprised of two sections. Section 1 provides an overview of the Cost/Schedule Control Systems Criteria (C/SCSC) to expose the reader to concepts relevant to this study. The second section is an in-depth review of the various criteria for constructing high-integrity graphs, as well as determining the amount of distortion within a graph. In addition to researching investigative question 2, this was done so the reader would have a good understanding of the experimental portion of this thesis, which tests how distorted and/or misleading graphs of CPR data can affect decision makers.

Section 1--C/SCSC Overview

Program managers are under increasing pressure to deliver products within cost and schedule limits. With the constant increase in technology, products are becoming increasingly sophisticated and costly. A good example is the B-2 bomber, costing approximately $570 million per copy (2:19). To keep these projects under control, managers utilize management control systems to monitor the cost, schedule, and technical status of their projects. Some management control systems utilize the Cost/Schedule Control System Criteria (C/SCSC). The Air Force, as well as all other Department of Defense components, Department of Energy, Department of Transportation, National Aeronautics and Space Administration, and the Federal Aviation

11

Administration, are using this sytem to manage selected contracts. Additionally, some foreign governments have adopted similar criteria (3:xi).

For government contractors working on a project in which the use of the C/SCSC has been mandated, cost and schedule information is usually reported on a monthly basis. On selected major weapon system acquisitions, the Cost Performance Report (CPR) is used and C/SCSC are usually applied. For programs too small for the full application of the C/SCSC, the Cost/Schedule Status Report (C/SSR) is used. The C/SSR is basically a scaled down version of the CPR.

> The concepts behind the C/SCSC are 1) the integration of management control systems such as planning, work scheduling, budgeting, authorizing work, cost accounting, and organizing; 2) periodic, objective statusing of planned work performed; and 3) associating budget with the work performed to produce an earned value. The C/SCSC in general and these three concepts in particular will provide a framework for pricing, planning, budgeting, and controlling functional organizations and programs of all sizes. They will ensure that sufficient data are available for detailed variance analysis and early problem detection (4:17).

Much planning needs to be accomplished prior to the application of a management control system utilizing C/SCSC. The starting point for planning is the Work Breakdown Structure (WBS), which completely defines the program and displays the relation of various subsystems to each other. The WBS breaks the work down into discrete "work packages." DOD Directive 5010.20 defines a WBS as follows:

A work breakdown structure is a product-oriented family tree composed of hardware, services and data which result from project engineering efforts during the development and production of a defense materiel item, and which completely defines the project/program. A WBS displays and defines the product(s) to be developed or produced and relates the elements of work to be accomplished to each other and to the end product (3:79).

Figure 6 contains an example of a WBS.

A "cost account" is comprised of at least one work package. The term "cost account" is used to describe the natural control point formed by the intersection of a given organization's functional responsibility with the effort required under a given WBS element (4:34). "Work packages are detailed short-span jobs, or material items, identified by the contractor for accomplishing work required to complete a contract (3:94)." Clearly, detailed planning must be accomplished to determine both the amount of effort required and the amount of money to be spent accomplishing the effort for each work package. Figure 7 relates the various cost accounts to an organizational structure.

The C/SCSC require that the contractor must be able to identify variances at the cost account level. However, the government must be provided with summarized data—summarized both for the WBS identified in the contract and for the functional organizations within the plant. Normally, the contractor is required to summarize progress at the third level of the WBS for reporting to the government (6:9). By reporting at a summary level, small variances, which always occur, tend to compensate for each other and cancel out.

13
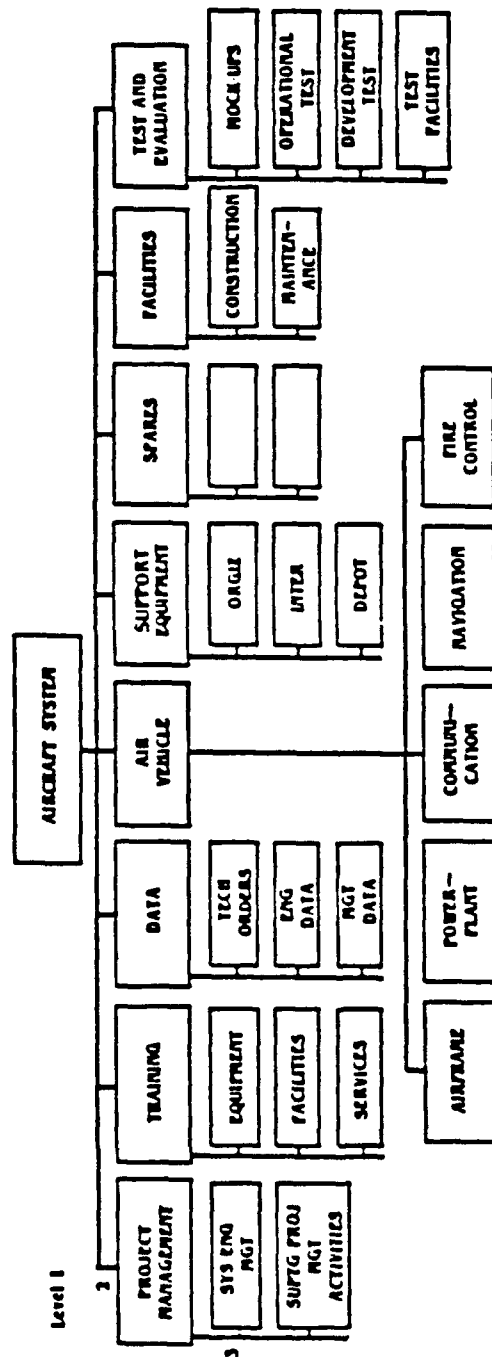
Figure 6. Work Breakdown Structure (Reprinted from 3:80)

14
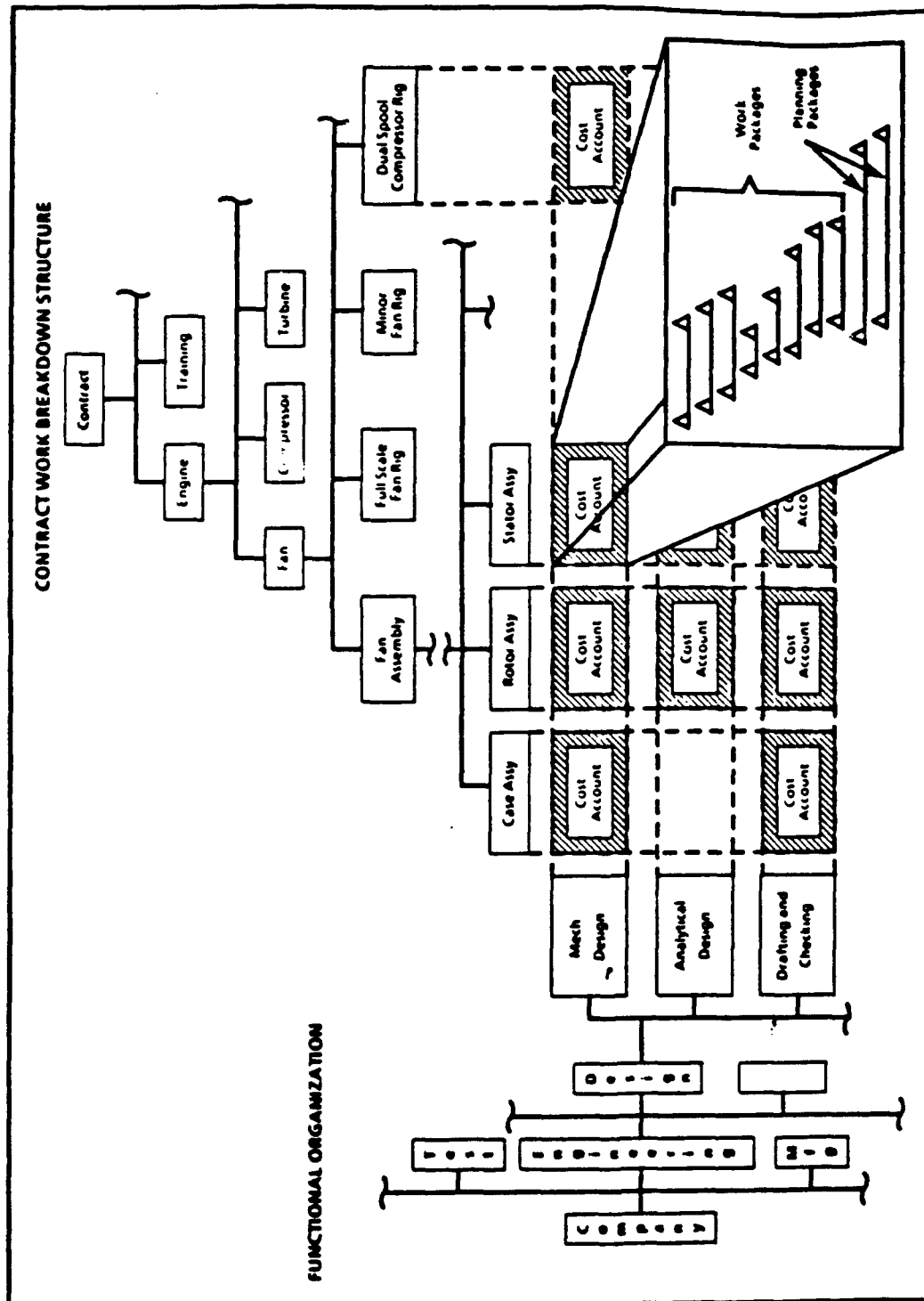
Figure 7. Integration of the WBS and the Organization
Structure (Reprinted from 5:35)

15

Only large variances worthy of management's attention will remain.

Managers must not only keep programs within cost, but also within schedule. Again, detailed planning must be done to determine the length of time each work package requires. Additionally, these schedules must be integrated to support the overall program schedule.

Clearly, the amount of planning for a project can become very involved, particularly for very large efforts. Why is all of this up-front work necessary? One answer is performance measurement. A baseline needs to be created to compare program cost and schedule performance against. Variance analysis is another reason for planning a project like this. Cost overruns, for example, are readily visible when a C/S^SC compliant management control system is being used correctly. Some techniques on how a manager can measure the performance of his program through the use of the C/SCSC shall be discussed.

First, each work package has a budget allocated for its completion. The Budgeted Cost of Work Scheduled (BCWS) is the budgeted cost for the work scheduled for a given period (4:33). This is the amount planned for the completion of a specific work package. As that work package is completed, the Budgeted Cost of Work Performed (BCWP) is computed. The BCWP is tne budgeted cost for the work actually performed during a given period. The Actual Cost of Work Performed (ACWP) is simply the actual cost for the work that was

16

performed (4:33). From these key data elements, a work package's cost performance can be easily computed:

$$\text{Cost Variance (CV)} = \text{BCWP} - \text{ACWP} \qquad (1)$$

Also, the work package's schedule performance may be computed:

$$\text{Schedule Variance (SV)} = \text{BCWP} - \text{BCWS} \qquad (2)$$

Cost and schedule variances of zero indicate the work package is right on cost and schedule goals. Negative variances indicate the work package is behind cost and/or schedule. Various combinations are possible, such as behind schedule and ahead of cost. The overall program's cost and schedule performance are the cumulative totals of the individual cost and schedule variances for the given period.

Finally, managers are interested in what the completed program will cost. A commonly used Estimate at Completion (EAC) projection technique is (6:21):

$$\text{EAC} = \frac{\text{BAC}}{\text{CPI}_{cum}} \qquad (3)$$

where

      BAC = Budget at Completion, the total budget for all authorized work.

      CPI_{cum} = Cost Performance Index, a measure of the cost effiency with which work has been accomplished.

$$= \frac{\text{Cumulative BCWP}}{\text{Cumulative ACWP}} \qquad (4)$$

17

Often, managers will keep track of the BCWS, BCWP, ACWP, and EAC by constructing a Cumulative Contract Status Graph (Figure 8). At any point prior to NOW, the cost and schedule variances are readily apparent. The variances also may be graphed separately; one example is a Current Month Variance Trend Graph. (There are many other types of graphs routinely used to portray CPR data).

Given an adequate management control system (i.e., C/SCSC-compliant), managers can more easily keep very large programs within cost and schedule constraints by identifying variances early. The CPR format also identifies which functional areas may be experiencing trouble which may adversely affect the program. Timely, periodic reporting and variance analysis are the keys to effective program management.

Section 2--Criteria for High-Integrity Graphs

Many authors have written about what constitutes a "good" graph. The first standards for good graphsmanship were proposed by the Joint Committee on Standards for Graphic Representation in 1915. Over the years, many other guidelines, or criteria, have been developed, some of which are expansions or modernizations to the Joint Committee's, and some of which are new ideas. In fact, some of the authors provide conflicting criteria. Table 1 contains a matrix of high-integrity graph criteria cross referenced to the authors who advocate them. A "X" indicates that a
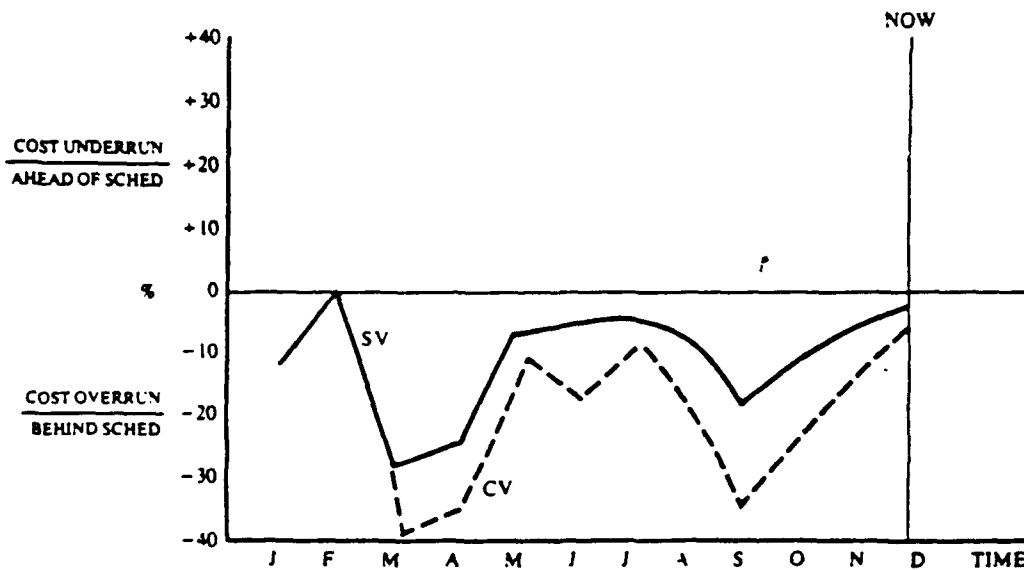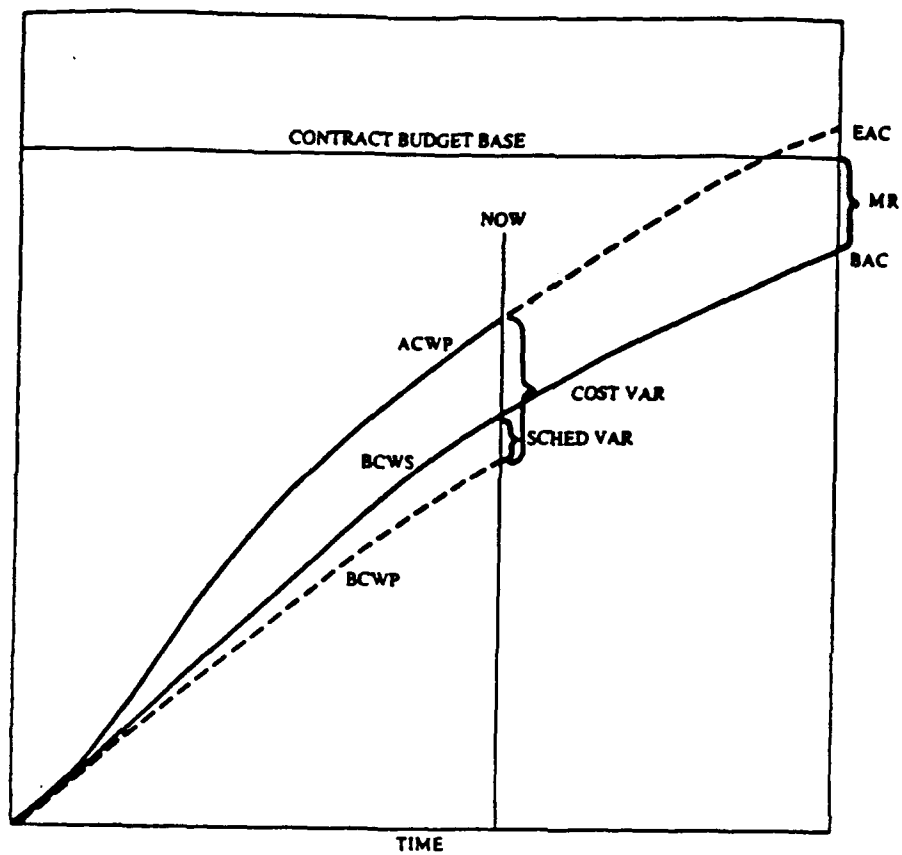
Figure 8. Cumulative Contract Status Graph (top) and
Current Month Variance Trend Graph (bottom)
(Reprinted from 4:17-18)

19

particular author endorses the use of some criterion, while an "O" indicates the author argues against its use. This table attempts to capture the high-integrity criteria with all who advocate their use in an easy-to-use format. In constructing this table, the exact words of some authors in stating some of the criteria may have been changed slightly to fit what other authors wrote. As such, the words may be different from what a specific author said, but the principle remains intact. Table 2 describes various style guides for creating well-designed graphics. Table 2 differs from Table 1 in that Table 2 does not address high-integrity graphical criteria, but rather formatting and procedural techniques.

Graphs should portray what the underlying data are implying. Frequently, readers may be overwhelmed by the way the graph presents the data, and retain a different impression than what the data suggests. Simplicity seems to be the key behind many criteria. "It is bad charting to tell too much on one chart. Anything that must be studied for its meaning is not good for popular presentation (7:176)." An overly "busy" or distorted chart may disinterest the reader, leaving the story of the data untold. "A chart frequently defeats its own purpose by leaving the viewer to unravel the total picture into significant components (17:47)."

Table 1. Criteria for Creating High-Integrity Graphics
With the Authors who Advocate Their Use

CRITERIA FOR CREATING HIGH INTEGRITY GRAPHICS	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Charts with an arithmetic scale should begin at the zero base line in order to show the true variation in movements.	X	X	X	X	X	X	X	X	X	X	X	X	X	O
2. Use multiple scales cautiously.		X		X							X			
3. The dependent axis should employ a simple arithmetic scale.						X		X						
4. Do not extend the scale much beyond the highest or lowest points on the graph.		X		X							X	X		X
5. If multiple curves are shown, the same unit scale must be used for correct comparison.												X		
6. Use labels to defeat graphical distortion and ambiguity.	X		X	X					X			X		
7. Represent quantities by linear magnitudes as areas or volumes may be misinterpreted.	X			X	X	X			X	X				
8. For area graphs, the more irregular strata should be placed near the top.														
9. Time scale divisions must be equal.				X		X				X	X	X		
10. Keep your charts simple to add to clarity.			X								X	X	X	

	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Table 1 Continued

AUTHORS

CRITERIA FOR CREATING HIGH INTEGRITY GRAPHICS	1	2	3	4	5	6	7	8	9	10	11	12	13	14
11. The horizontal scale should usually be read from left to right; the vertical scale from bottom to top.					X									
12. The general arrangement of a graph should proceed from left to right.					X									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14

AUTHORS	YEAR	BIBLIOGRAPHICAL REFERENCE NUMBER
1. Tufte	1983	8
2. Taylor and Anderson	1986	9
3. Cox	1978	10
4. Schmid	1954	11
5. Joint Committee on Standards for Graphic Representation	1915	12
6. MacGregor	1979	13
7. Steinbart	1986	14
8. Johnson, Rice, and Roomich	1980	15
9. Spear	1969	7
10. Auger	1979	16
11. Rogers	1961	17
12. American Society of Mechanical Engineers	1979	18
13. Lefferts	1981	19
14. Cleveland	1985	20

Table 2. Style Guides for Creating Good Graphics

AUTHORS

"STYLE GUIDES" FOR CREATING GOOD CHARTS	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Scale breaks should be used for false origins.			X	X	X							X		
2. Graphics must not quote data out of context.	X													
3. Oblong shaped grids are preferable to square grids. Good standard proportions are two to three and three to four.			X									X		
4. The zero lines should be sharply distinguished.			X	X					X					
5. The curve lines should be distinguished from the grid ruling.					X				X		X	X		
6. Try to include in the diagram the numerical data.			X											
7. If the data is not included, give the data in tabular form accompanying the diagram.			X								X	X		
8. When shading, shade from the zero line to the curve.									X					
9. Vertical or horizontal shadings are not recommended.									X			X	X	
10. Patterned shadings should be of good contrast.									X			X		

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |

23

Table 2 Continued

AUTHORS

"STYLE GUIDES" FOR CREATING GOOD CHARTS	1	2	3	4	5	6	7	8	9	10	11	12	13	14
11. Legends should make diagrams nearly self-explanatory.			X											
12. Scales should be such that linear relations are roughly 45 degrees to the x-axis.			X											
13. For column charts, the columns should be the same width; spacing between is one-half the column width.						X			X					
14. Arrange columns systematically.			X						X		X			
15. When a large part of the grid is unnecessary, break the grid but retain the zero line.						X								
16. Eliminate all grid lines but those essential for easy reading.	X		X	X							X	X		
17. On multiple curve graphs, each curve should be the same width.						X					X	O		
18. If irregularities occur in the time sequence, include spaces for the missing columns.									X					
19. Avoid broken scales which give innaccurate impressions.				X						X		X		X
20. Standardized units of monetary measurements are better than nominal units.	X													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Table 2 Continued

"STYLE GUIDES" FOR CREATING GOOD CHARTS	1	2	3	4	5	6	7	8	9	10	11	12	13	14
21. For most line charts the maximum number of plotted lines should not exceed five; three or fewer is the ideal number.													X	
22. The simplest curve patterns are usually the most effective. A solid line is most useful.												X		
23. Keep your charts as simple as possible to add to clarity.											X			X
24. D⁻ not overdo the number of tickmarks.														X
	1	2	3	4	5	6	7	8	9	10	11	12	13	14

While there are no concrete rules that must be followed while constructing graphs, several researchers have expressed the need for such. According to Feinberg,

> A rational set of graphical standards should be based on a theory for graphic representation. Alas, we have no such theory, and the prospects for its development remain dim. Yet, it is easy to come up with a simple set of suggestions that would improve the clarity of most graphs (21:166).

The American Society of Mechanical Engineers counters this argument by suggesting, "There is no fixed formula for chart design; each chart must be 'hand tailored' to fit the needs of the problem" (18:ix). However, Fienberg is correct in saying most graphs would have improved clarity if they were constructed in accordance with a set of standards. Tables 1 and 2 are an attempt to provide these standards.

Not all authors agree on every criterion. Cleveland argues, in retaliation to criterion 1 of Table 1, "that the compulsion to include zero on a scale has ruined many graphs (20:78)." Refering back to Figure 1, Cleveland would insist the graph is a waste of space because the resolution is so poor. A better graph, in his judgment, might have the closing figure scale start at 2,000 and end at 3,000. According to Cleveland, "this new graph would convey more quantitative information in the sense the reader would be able to discern variations in the data more readily (20:79)." However, Table 1 shows 13 authors who would find this new graph misleading. Other cases of conflict exist

between authors, with the potential of confusing any student of the material.

Some authors have gone further than proposing good criteria to follow when constructing graphs. One has even developed a method to measure the distortion found in misleading graphs. This method is the Tufte Lie Factor. According to Tufte, "the representation of numbers, as physically measured on the surface of the graph itself, should be directly proportional to the numerical quantities represented (8:56)." The Lie Factor model measures the amount of misrepresentation in graphs by this formula:

$$\text{Lie Factor} = \frac{\text{Size of Effect Shown in Graphic}}{\text{Size of Effect in Data}} \qquad (5)$$

According to Tufte,

> If the lie factor is equal to one, then the graphic might be doing a reasonable job of accurately representing the underlying numbers. Lie factors greater than 1.05 or less than .95 indicate substantial distortion, far beyond minor discrepancies in plotting (8:57).

Appendix A contains examples of distorted graphs in widely read publications, and the amount of misrepresentation as computed by the lie factor formula. Clearly, distortion of graphs is very common, and probably, unintentional. The informed reader should beware of distortion to avoid being misled.

Given that some of these high integrity graph criteria have been in existence for decades, it is difficult to see

27

why prestigious publications like the Wall Street Journal routinely include misleading and/or distorted graphics in virtually every issue. Of course, they are in the business of selling newspapers to make a profit. Thus, their graphs may be designed to grasp a reader's attention by graphical style, with graphical integrity being a secondary concern. Although the Wall Street Journal does not go overboard with ornate styling, it tends to vary the scales used tremendously. The effect looks like a wildly fluctuating entity. Perhaps readers are drawn to such graphs rather than relatively stable ones. The reader must know if information being presented is distorted, and must be able to filter out the true meaning.

Summary

The literature review has confirmed the existence of criteria, which, when followed, will help the graph maker produce high-integrity graphs free from distortion and potentially misleading effects. In addition, many "style guides" have been developed to enable graph makers to portray information more clearly and effectively. Thus, investigative question 2 has been answered: there are criteria, principles, or guidelines of graphical excellence.

Most sources of the criteria were from the academic and business communities. No Department of Defense sources were found. This is not disconcerting, as most of the graphs used in the Air Force acquisition community, in the author's

opinion, do not differ greatly from those in use elsewhere.
Bar graphs, line graphs, area graphs, etc., are very
prevalent in the Air Force and the defense community. The
criteria researched in the literature review apply just as
well to these graphs as they do to non-defense related
activity graphs.

Now that criteria for creating high-integrity graphs
have been documented, investigative question 4, which
examines if graphs of CPR data which violate the criteria
can mislead decision makers, may be researched. Chapter
III, describes the experiment and statistical analysis
methods used to answer this question.

III. Methodology

The objectives of this thesis are two-fold: determine
what graphs of CPR data are used to formulate managerial
decisions, and determine if distorted graphs of CPR data can
lead to faulty decisions. The investigative questions are
as follows:

1. What kinds of graphs of CPR data are being used?

2. Are there criteria, principles, cr guidelines of
graphical excellence?

3. Are program managers/analysts creating graphs which
violate the criteria?

4. Can graphs of CPR data which violate the criteria
mislead decision makers?

Investigative questions 1 and 3 were researched by
collecting samples of graphs from throughout Air Force
Systems Command (AFSC). These graphs were requested through
Brigadier General John M. Nauseef, AFSC DCS/Comptroller. A
copy of General Nauseef's letter requesting samples of
graphs constructed by SPOs throughout AFSC is contained in
Appendix B. Investigative question 2 was researched through
a literature review, and the results are contained in
Chapter II, Tables 1 and 2.

To answer research question 4, an experiment was
designed to test AFIT Professional Continuing Education
(PCE) students. This experiment attempted to determine if
Air Force decision makers could be misled by graphs chat
violate the criteria for constructing high-integrity

graphics. The experimental design, all graphs used in the experiment, and the statistical analysis techniques shall now be discussed.

Experimental Design

To test if Air Force decision makers can be misled by graphs that violate high-integrity criteria, experiments were conducted on AFIT PCE students. These students, predominantly from the cost analysis/program control careers, were chosen to participate in the experiment because they represent likely decision makers in weapon systems acquisition programs throughout the Air Force. In cases where the cost analyst would report to a program manager, it is likely the manager would request the cost analyst's opinions and recommendations. A secondary reason for including these PCE students in the experiment was their availability.

The experiment used a Pretest-Posttest Control Group design (22:13):

$$R \ O_1 \ X \ O_2$$

$$R \ O_3 \ \ \ O_4$$

First, the students were divided into two groups, the $O_1 \ X \ O_2$ (experimental) group, and the $O_3 \ \ \ O_4$ (control) group. The "R" preceding each group indicates a random division of tne population was used to separate test subjects into the two groups. However, each subject's experience in the C/SCSC field was considered in the group

assignment procedure to ensure the two groups were as equal in terms of experience as possible. The "X" in the experimental group indicates where the treatment was applied. This design calls for exposing both groups to a pre-test first (O_1 for the experimental group and O_3 for the control group). Next, a posttest is applied (O_2 and O_4 for the experimental and control groups, respectively). The X indicates the experimental group's posttest had the treatment applied to it; in this case, the treatment was the misleading graphics in the posttest.

According to Campbell and Stanley, the typical threats to internal invalidity are (22:5):

> 1) history, the specific events occurring between the first and second measurement in addition to the experimental variable;
>
> 2) maturation, or processes within the respondents operating as a function of the passage of time, including growing older, hungrier, more tired, etc;
>
> 3) testing, which describes the effects of taking a test upon the scores of the second testing;
>
> 4) instrumentation, in which changes in the calibration of a measuring instrument or changes in the observers or scorers used may produce changes in the obtained measurements;
>
> 5) regression, where groups have been selected on the basis of their extreme scores;
>
> 6) selection, or placement of subjects in different groups;
>
> 7) mortality, or differential loss of respondents from the comparison group;
>
> 8) selection-maturation, which might be mistaken for the effect of the experimental variable (in this case, misleading graphics).

All of these threats are controlled using the Pretest-Posttest Control Group Design. History is controlled insofar as general historical events that might have produced an $O_1 - O_2$ difference would also produce an $O_3 - O_4$ difference (22:13). Since the posttest immediately followed the pretest, history can further be ruled out as a threat to internal validity. "Maturation and testing are controlled in that they should be manifested equally in the experimental and control groups (22:14)." To combat the effects of maturation, the first trial was conducted immediately after a lunch break, while the second trial was run as the first activity of the day. Theoretically, the students should have been well rested, not hungry, and not yet eager to complete their work for the day. Testing is of no concern here, as there were no repeat subjects. No one took the experiment twice. Since a printed test is being administered, instrumentation is easily controlled. Regression and selection are ruled out as causes of invalidity because no placement tests were used to assign subjects to groups; rather, they were randomly assigned. Mortality is not a concern, as all cases were used in the analysis.

The factors jeopardizing external validity are (22:5):

> 1) the reactive or interaction effect of testing, in which a pretest might increase or decrease the respondent's sensitivity or responsiveness to the experimental variable and thus make the results obtained for a pretested population unrepresentative of the effects of the experimental variable for the

33

unpretested universe from which the experimental respondents were selected;

2) the interaction effects of selection biases and the experimental variable;

3) the reactive effects of experimental arrangements, which would preclude generalization about the effect of the experimental variable upon persons being exposed to it in nonexperimental settings;

4) multiple-treatment inference, likely to occur whenever multiple treatments are applied to the same respondents, because the effects of prior treatments are not usually erasable.

According to Campbell and Stanley, the only factor that is a major weakness of this design is the first one listed, the reactive or interaction effect of testing (22:8). More will be said about this in Chapter 4, Analysis and Findings.

To test for differences between pretest and posttest scores for both groups, the following experimental hypotheses were developed:

Ho: $(O_1 - O_2) - (O_3 - O_4) = 0$; the graphs were not misleading.

Ha: $(O_1 - O_2) - (O_3 - O_4) > 0$; the graphs were misleading.

For this experiment, the null hypothesis shall be rejected at alpha = .05. These hypotheses measure the change in both group's scores from pretest to posttest. Since the control group received charts developed in accordance with the high-integrity criteria, one would not expect a significant difference in scores from pretest to posttest. In other words, the quantity $(O_3 - O_4)$ should equal zero. On the

34

other hand, the quantity $(O_1 - O_2)$ may be expected to be quite large, as the pretest consisted of good charts, while the posttest was comprised of misleading charts. If this is true, at some level of statistical significance, the null hypothesis may be rejected.

Constructing and Conducting the Experiment

The first step in constructing the experiment was to identify graphs commonly used in cost analysis, specifically graphs used in a management control system utilizing the C/SCSC. Two sources were used to find graphs suitable for this experiment: the survey returns used to answer research questions 1 and 3, and AFSCP 173-4, Guide to Analysis of Contractor Cost Data. From these sources, six types of charts were chosen.

Secondly, Table 1 (from Chapter II) was modified. Instead of cross-referencing criteria for high integrity graphics to authors, they were cross-referenced to the six types of charts chosen for the experiment. The result is Table 3, which identifies which criterion may be violated to make each particular chart misleading. From this table, it was easy to construct graphs which violated at least one of the criteria, resulting in a misleading graph. Each "X" in Table 3 indicates that the charts selected for use in the experiment may be constructed in a misleading manner by violating the corresponding criterion. The boxed "X" reflects which criterion was used to create the misleading

35

Table 3. Commonly Used Charts in CPR Analysis Cross-Referenced
to Criteria Which, When Violated, May Cause Graphical
Distortion or Prove Misleading.

CRITERIA FOR CREATING HIGH INTEGRITY GRAPHICS	COMMONLY USED CHARTS IN CPR ANALYSIS					
	1	2	3	4	5	6
1. Charts with an arithmetic scale should begin at the zero base line in order to show the true variation in movements.		X	X	X	X	X
2. Use multiple scales cautiously.						
3. The dependent axis should employ a simple arithmetic scale.	X	X	X	X	X	
4. Do not extend the scale much beyond the highest or lowest points on the graph.	X	X	X	X	X	
5. If multiple curves are shown, the same unit scale must be used for correct comparison.				X	X	
6. Use labels to defeat graphical distortion and ambiguity.	X	X	X	X	X	X
7. Represent quantities by linear magnitudes as areas or volumes may be misinterpreted.	X		X			
8. For area graphs, the more irregular strata should be placed near the top.		X				
9. Time scale divisions must be equal.		X	X	X	X	
10. Keep your charts simple to add to clarity.	X	X	X	X	X	X
	1	2	3	4	5	6

Table 3 Continued

CRITERIA FOR CREATING HIGH-INTEGRITY GRAPHICS	COMMONLY USED CHARTS IN CPR ANALYSIS					
	1	2	3	4	5	6
11. The horizontal scale should usually be read from left to right; the vertical scale from bottom to top.	[X]	X	X	[X]	X	X
12. The general arrangement of a graph should proceed from left to right.	X	X	X	X	[X]	
	1	2	3	4	5	6

CHARTS

1. B-3 Subcontractor Variances (Cumulative)--rotated bar graph.

2. Annual Direct and Indirect Costs--area graph.

3. Monthly Cost Data--volume graph.

4. Performance Index Trends--line graph.

5. MR Comparison to CV and SV--line graph.

6. Cumulative Cost and Schedule Variance Percentage--"Bull's Eye graph."

graph used in the experiment. Many different criteria may be used to create a misleading graphic. However, for each chart used in the experiment, only one criterion was violated. The pairing of a chart to a criterion was based on the author's judgment.

During the actual experiment, the control group members received a package containing high-integrity graphs, along with the pretest charts. In addition to the same pretest charts that the control group received, the experimental group members were given a series of misleading graphs. By comparing the results of the two groups, a test can be conducted to see if the manner in which graphs are constructed has an effect on how a user interprets the information contained in the graph. Each pair of graphs (high-integrity and misleading) were derived from identical data. The only difference between them is the way in which they were graphed.

Each subject received one test package of twelve graphs with one conclusion corresponding to each graph (copies of the pretest package along with the control group and experimental group posttest packages are contained in Appendix C). The subjects were told to keep the packages face-down until they were told to read the cover letter. The cover letter briefly mentioned how program managers and cost analysts were increasingly reliant on graphs to analyze contractor cost data, and how this could be done very quickly and easily with the availability of desktop

computers and spreadsheet software. Great care was taken to avoid any mention of misleading graphics and control/experimental groups so as not to bias the experiment. The subjects were also told that each graph was independent of the others; i.e., the information in any one graph had nothing to do with the information in any other graph. They were also asked to refrain from looking at other subject's packages. The sequence of graphs in each test package was randomized to discourage one from trying to see how another subject answered the conclusion. Once everyone was ready, the signal to begin was given.

The subjects were given thirty seconds to review each graph, and then were told to flip the page over, read the conclusion, and indicate whether or not they agreed with it by circling the appropriate response ("Agree" or "Disagree"). Fifteen seconds were allowed for this. These relatively short amounts of time were given to simulate the amount of time a busy program manager or cost analyst would have to review graphs like these. The correct answers to the posttest were split evenly between "Agree" and "Disagree" to avoid the possibility of a subject answering all conclusions correctly or incorrectly by merely circling the same answer out of boredom or lack of interest in the experiment. Copies of the actual charts used, as well as the steps undertaken in their preparation, shall now be described. Each chart on the following pages (Figures 9-14) is shown as a high-integrity graph at the top of the

figure, and a misleading graph at the bottom of the figure.
The conclusion pertaining to each graph is also included.
In summary, each subject was exposed to high-integrity OR
misleading charts in the posttest. They did not see both
versions of each graph, as Figures 9 - 14 may imply.

Chart 1. Figure 9 contains Chart 1, titled "B-3
Subcontractor Variances (Cumulative)." The top figure shows
the chart constructed in a high integrity manner; i.e., free
from distortion or misleading effects. To transform this
graph into a misleading graph, criterion 11 (The horizontal
scale should usualy be read from left to right; the vertical
scale from bottom to top.) was chosen from Table 3. Note
the horizontal scale is reversed in the bottom figure,
creating the impression that only one subcontractor is
behind schedule. Assuming that a decision-maker would
interpret a bar to the left of zero as negative, the
unsuspecting decision maker could be fooled by the chart on
the bottom of figure 9.

Chart 2. Chart 2, "Annual Direct and Indirect
Costs", is presented as Figure 10. Criterion 8 (For area
graphs, the more irregular strata should be placed near the
top.) was chosen to create a misleading graph. The rule to
remember when reading area graphs is to measure each stratum
from the one immediately below it, not from zero. The top
graph shows clearly the indirect costs have been constant
throughout the time span. The graph at the bottom also
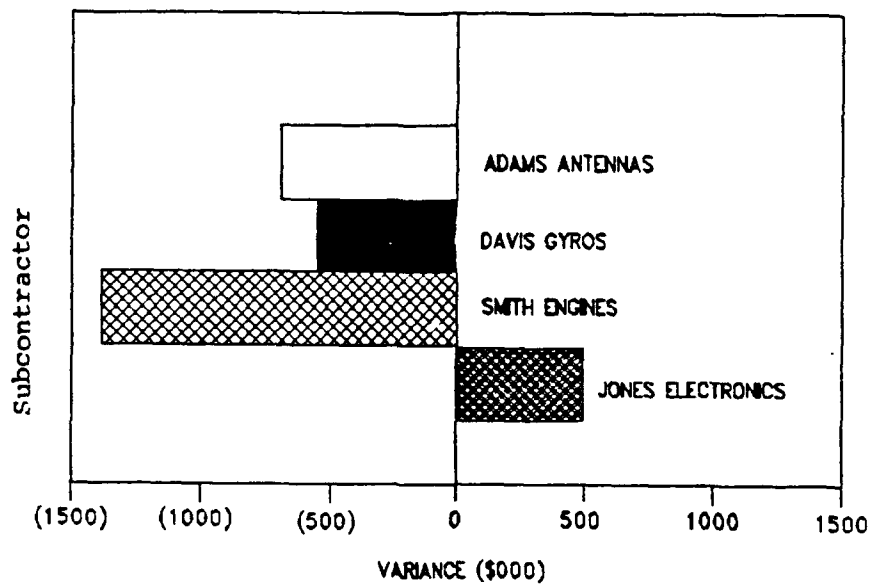
40

Conclusion: Of the subcontractors shown, Jones Flectronics
has the largest negative schedule variance.

DISAGRFE   AGRFE
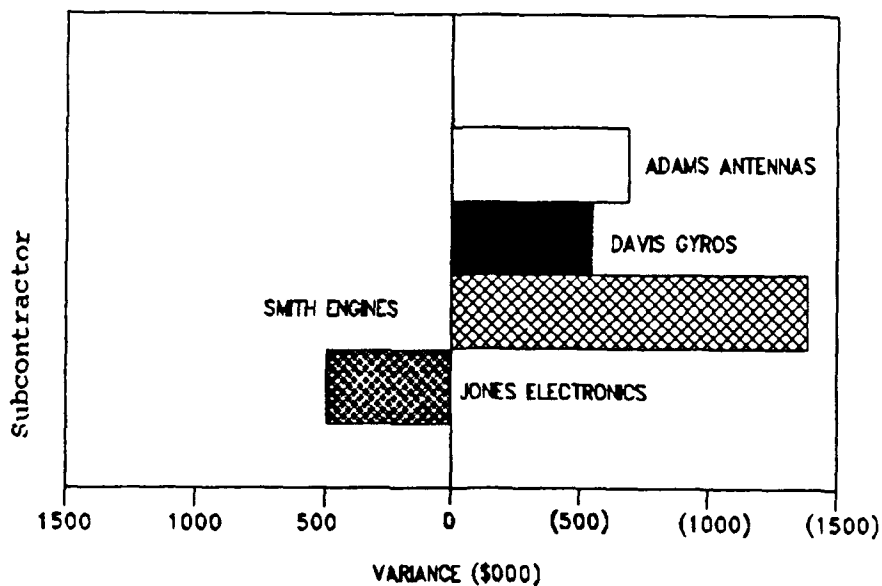
Figure 9. Chart 1.

41
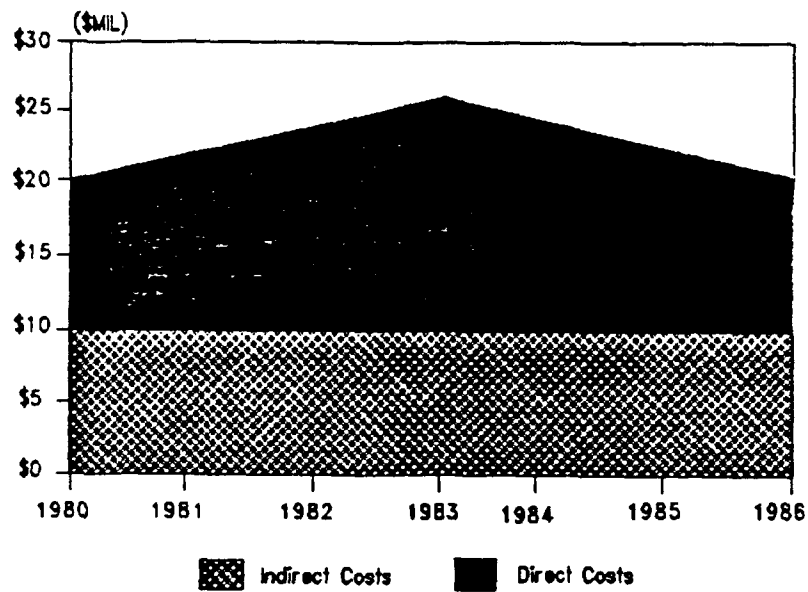
Conclusion: Indirect costs have been the same each year.

DISAGREE   AGREE
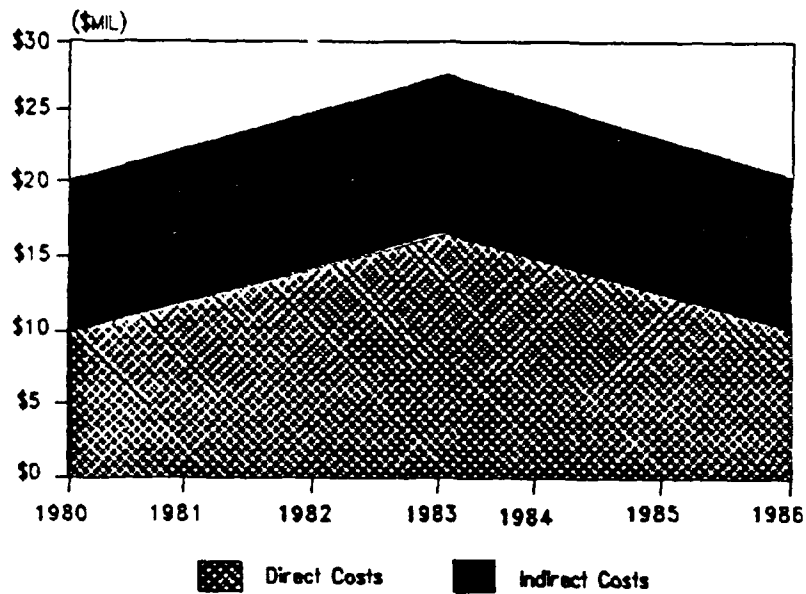
Figure 10.   Chart 2

reflects this, but in a manner that may be read as indirect costs are decreasing.

Chart 3. Figure 11 contains chart 3, titled "Monthly Cost Data." Here, the different entities are represented by linear magnitudes, as shown at the top. By violating criterion 7, substituting volumes in place of linear magnitudes, a misleading graph can be created, as shown at the bottom. In theory, "the number of information carrying dimensions depicted should not exceed the number of dimensions in the data (8:77). Only the front faces of the volumes in the bottom graph carry information; the magnitude of the volume is meaningless. Thus, a high-integrity graph picturing a constant ACWP has been changed to portray a steadily decreasing ACWP.

Chart 4. Criteria 11 was again used to create chart 4, "Performance Index Trends", shown in Figure 12. For this graph, an unfavorable trend in the schedule performance and cost performance indexes, shown at top, has been made to appear favorable by increasing over time. Note the vertical scale has been reversed in the bottom graph, creating a totally different image from the high-integrity version.

Chart 5. Chart 5, "MR Comparison to CV and SV", is shown in the high-integrity and the misleading versions in Figure 13. Criterion 12 (The general arrangement of a graph should proceed from left to right.) was employed to create
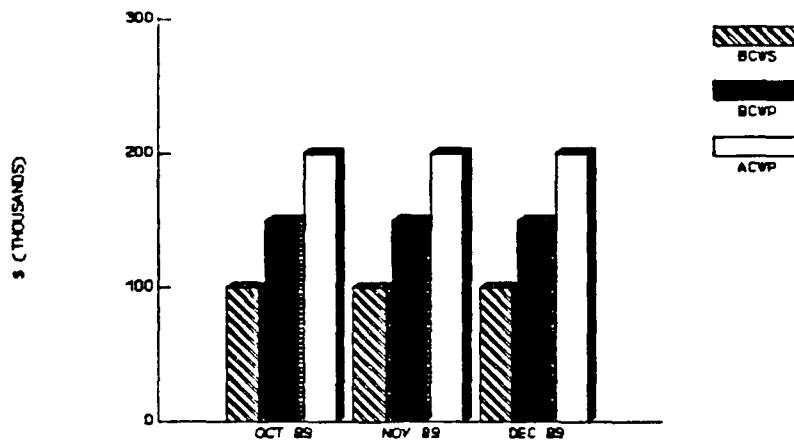
43

MONTHLY COST DATA

Conclusion: The ACWP is steadily decreasing throughout the period.
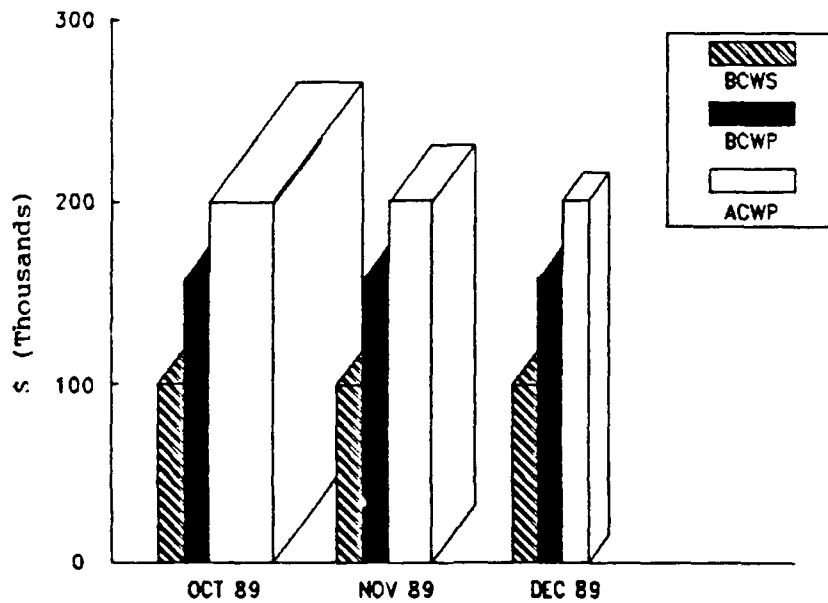
DISAGREE   AGREF

Figure 11.   Chart 3.

PERFORMANCE INDEX TRENDS
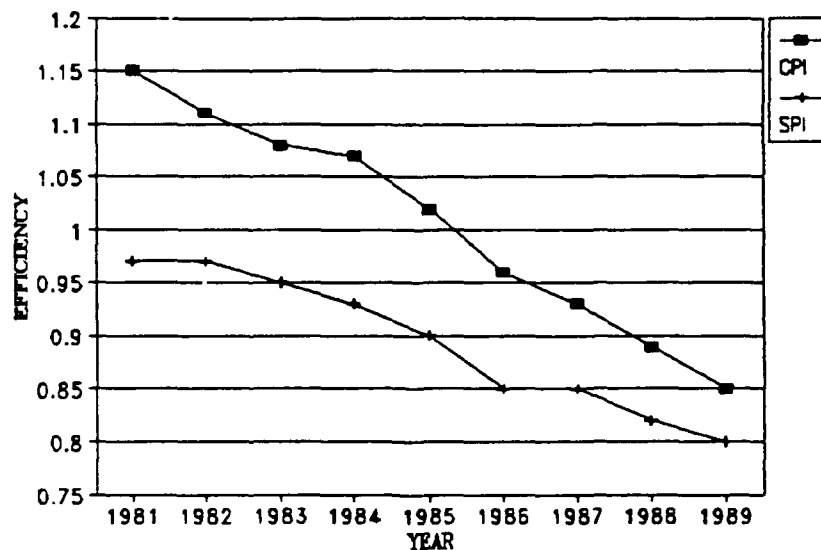
Conclusion: The CPI and SPI values in 1989 are better than those in 1981.

DISAGREE   AGREE

PERFORMANCE INDEX TRENDS
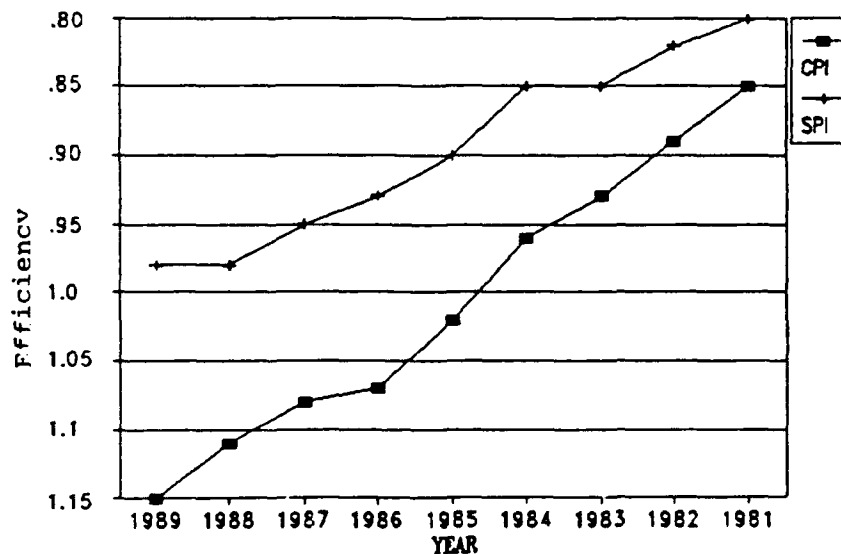
Figure 12. Chart 4.

MR COMPARISON
TO CV AND SV
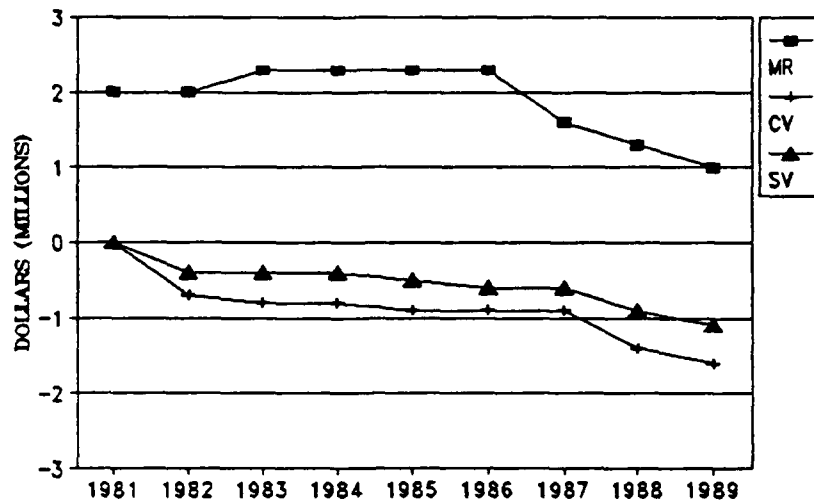
Conclusion: The cost variance and schedule variance reached their worst levels in 1989.

DISAGREE   AGREE

MR COMPARISON
TO CV AND SV
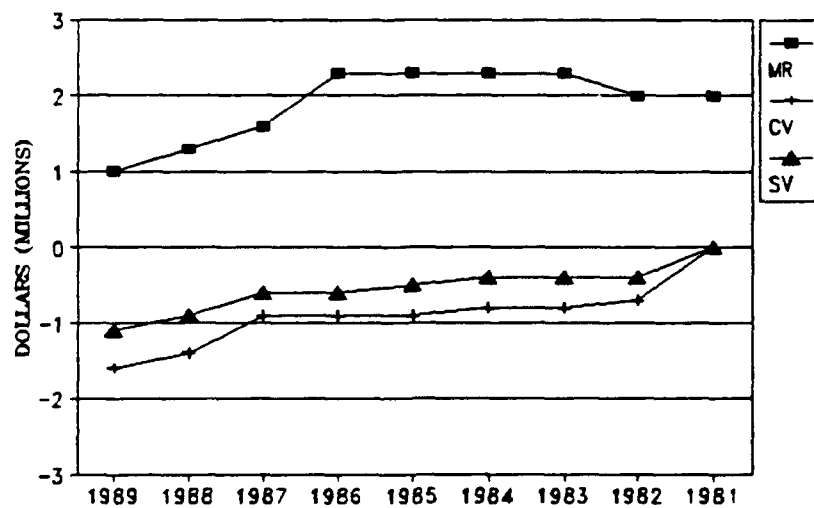
Figure 13. Chart 5.

the misleading graph. Reading the misleading graph on the bottom from right to left would convey the correct impression that both cost and schedule variances are growing worse, but people generally read charts and text from left to right. Reversing the time scale creates the false impression of favorable trends in both variances.

Chart 6. Labeling can be very useful in reading and understanding graphs. A good example is chart 6, "Cumulative Cost and Schedule Variance Percentage (6 Months)." As shown in Figure 14, incorrect labeling can possibly create different impressions on users. Criterion 6 (Use labels to defeat graphical distortion and ambiguity.) was used to derive a misleading version of the data from the high-integrity graph. Look closely at the first and third quadrants of the top graph. A program at point 1 would be ahead of schedule and underrunning costs (a good position to be in), while points 4, 5, and 6 would indicate the program was behind schedule and overrunning costs (which should be avoided). However, the misleading version at the bottom states point 1 is overrunning costs, while points 4, 5, and 6 are underrunning costs. Presumably, labeling this "Bullseye" chart's quadrants should help the user, but what if the labeling is wrong?

After all graphs were completed, the subjects were asked to look at a series of pages, with both high-integrity and misleading versions of the graphs on the same page. A

Cumulative Cost and Schedule Variance Percentage (6 Months)
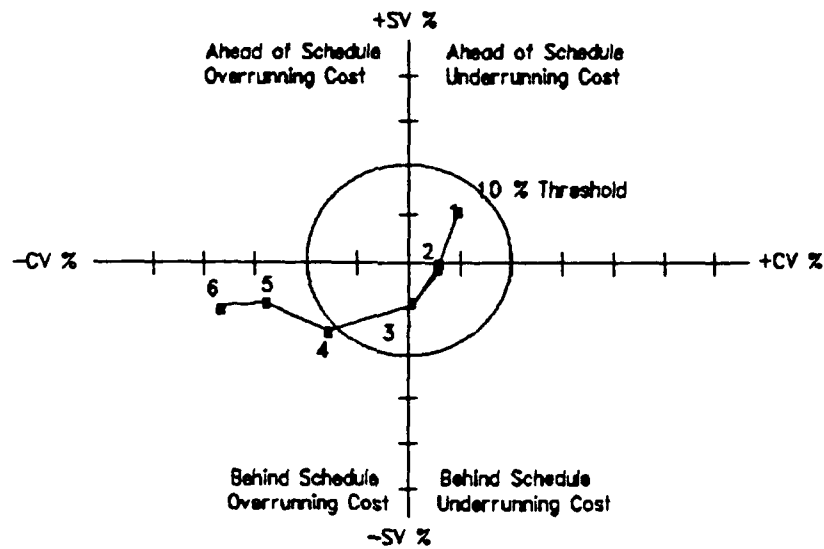
+SV %

Ahead of Schedule   Ahead of Schedule
Overrunning Cost   ┼   Underrunning Cost

10 % Threshold

-CV % ─────────────┼──────── 2 ┼─────────────── +CV %
                6   5
                         4   3

Behind Schedule   ┬   Behind Schedule
Overrunning Cost   │   Underrunning Cost

-SV %

Conclusion: At month 6, this program was behind schedule and
              overrunning cost.

DISAGREE   AGREE

Cumulative Cost and Schedule Variance Percentage (6 Months)

+SV %

Ahead of Schedule   Ahead of Schedule
Underrunning Cost   ┼   Overrunning Cost

10 % Threshold

                                    1
-CV % ─────────────┼──────── 2 ┼─────────────── +CV %
                6   5
                         4   3

Behind Schedule   ┬   Behind Schedule
Underrunning Cost │   Overrunning Cost

-SV %

Figure 14.   Chart 6.

48
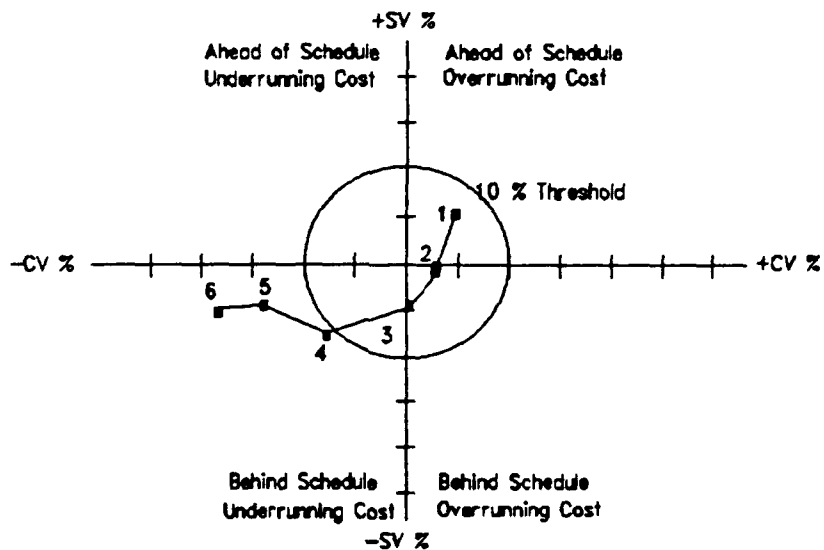
brief description of what was done to convert the good graph into the misleading graph was provided. The subjects were asked to write down any comments they had about the graph directly on the page. This information was then considered in making necessary revisions to improve the experimental instrument.

Statistical Analysis

To ease the task of tabulating the experimental results, Table 4 was developed. Since the statistical hypotheses are designed to test for differences between pretest and posttest scores for both groups, the table is divided into pretest and posttest sections. The pretest section is denoted by charts P1 through P6, while the posttest section is denoted simply by 1 through 6, indicating charts 1 through 6. Using dummy data for illustrative purposes, Ta. e 4 was completed by marking a "1" for a correct response, and a "0" for an incorrect response. For each student, the total number of correct answers in both pretest and posttest sections were computed. Additionally, the number of correct responses for each chart was also calculated.

A t-test was used to evaluate this data. The assumptions made for the analysis are: 1) the number of students is large; 2) the variances of each group are unknown; and 3) the variances are unequal. This test statistic was chosen for use because the number of students

49

Table 4. Data Tabulation Example

Experimental Group

Student	Q1 P1	P2	P3	P4	P5	P6	Tot	Q2 1	2	3	4	5	6	Tot	Delta	sigsq
1	1	1	0	1	1	1	5	0	0	1	0	0	1	2	3	0.871
2	1	1	1	1	1	1	6	1	0	0	0	0	0	1	5	1.138
3	1	0	1	1	1	1	5	0	0	0	1	0	0	1	4	0.004
4	1	1	1	1	1	1	6	1	0	0	0	1	0	2	4	0.004
5	1	1	0	1	1	1	5	1	0	1	1	0	0	3	2	3.738
6	1	1	1	1	1	1	6	0	0	0	0	0	1	1	5	1.138
7	1	1	1	0	1	1	5	1	0	0	0	0	0	1	4	0.004
8	0	1	1	1	1	1	5	0	1	0	0	0	0	1	4	0.004
9	1	1	1	1	1	1	6	1	1	0	0	0	0	2	4	0.004
10	1	0	1	1	1	1	5	0	0	0	0	0	0	0	5	1.138
11	1	1	1	1	1	0	5	0	1	0	0	0	0	1	4	0.004
12	1	1	1	0	1	1	5	0	0	0	0	1	0	1	4	0.004
13	1	0	1	1	1	1	5	1	1	0	0	0	0	2	3	0.871
14	1	1	1	1	1	1	6	0	0	0	1	0	1	2	4	0.004
15	1	0	1	1	1	1	5	1	0	0	0	0	0	1	4	0.004
Totals:	14	11	13	13	15	14	80	7	4	2	3	2	3	21	59	8.933

Mdelta = 3.933

Control Group

Student	Q3 P1	P2	P3	P4	P5	P6	Tot	Q4 1	2	3	4	5	6	Tot	Delta	sigsq
1	1	1	1	1	1	1	6	1	1	0	1	1	1	5	1	0.284
2	0	1	1	1	1	1	5	1	1	1	1	0	1	5	0	0.218
3	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.218
4	0	1	1	1	1	1	5	1	1	1	1	0	0	4	1	0.284
5	1	1	1	1	1	1	6	1	1	1	1	1	0	5	1	0.284
6	1	1	1	0	1	1	5	0	1	1	1	1	1	5	0	0.218
7	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.218
8	0	0	1	1	1	1	4	1	0	0	0	1	1	3	1	0.284
9	1	1	1	0	1	1	5	1	1	1	1	0	1	5	0	0.218
10	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.218
11	1	1	1	1	1	1	6	0	1	1	1	1	1	5	1	0.284
12	0	1	1	1	1	0	4	0	1	1	0	1	1	4	0	0.218
13	1	1	1	1	1	1	6	1	1	1	1	1	0	5	1	0.284
14	1	1	1	1	1	1	6	0	1	1	1	1	1	5	1	0.284
15	1	1	1	1	0	1	5	1	1	1	1	1	0	5	0	0.218
Totals:	11	14	15	13	14	14	81	0 11	14	13	13	12	11	74	7	3.733

Mdelta = 0.467

participating in the experiment was deemed large enough to preclude the use of a non-parametric test (i.e., greater than 30). However, the group variances were unknown and could not be presumed equal. Thus, the test takes the form of a "Case 3" test on the difference of two means (23:197). The statistical analysis approach, using the data from Table 4 for illustrative purposes, shall now be discussed.

First, the difference between pretest and posttest scores was tabulated for each student. Then, each individual variance was computed using the following formula (23:53):

$$\sigma^2 = \sum_{i=1}^{n} (x_i - u)^2 \tag{6}$$

The next step is to compute the sample variance for both control and experimental groups using this formula (23:105):

$$S_x^2 = \frac{\sum_{i=1}^{n} (x_i - x)^2}{n - 1} \tag{7}$$

The remaining step is to plug all variables into the Case 3 t-statistic equation (23:197):

$$T_f = \frac{x_1 - x_2 - (u_1 - u_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \tag{8}$$

51

where n_1 and n_2 are the sample sizes and S_1^2 and S_2^2 are the sample variances. The degrees of freedom associated with the random variable T is f where (23:198):

$$f = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\dfrac{(S_1^2/n_1)^2}{n_1 - 1} + \dfrac{(S_2^2/n_2)^2}{n_2 - 1}} \qquad (9)$$

Thus, T_f has a t-distribution with f degrees of freedom.

Using the Table 4 data, a t-statistic shall now be computed using formulas (6) through (9). Since the null hypothesis states there is no difference in the number of correct answers between the control and experimental groups, $u_1 - u_2 = 0$. Thus, formula (8) now can be written as:

$$T_f = \frac{x_1 - x_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \qquad (10)$$

Now, the sample variances may be computed utilizing formula (7):

$$S_{exp}^2 = \frac{8.933}{14} = .63807$$

$$S_{con}^2 = \frac{3.733}{14} = .26664$$

As shown in table 4, $x_{exp} = 3.9333$ and $x_{con} = .467$. Plugging all variables into formula (10),

$$T_f = \frac{3.9333 - .467}{\sqrt{\dfrac{.63807}{15} + \dfrac{.26664}{15}}} = 14.11423$$

52

where

$$f = \frac{(.63807/15 + .26664/15)^2}{\dfrac{(.63807/15)^2}{14} + \dfrac{(.26664/15)^2}{14}} = 23.96127$$

Results of Each Posttest Chart

The charts having the most misleading effect can be identified. Two-way contigency tables can easily compute the chi-squared values, from which a level of significance may be obtained. For example, the contigency table for chart 1, using Table 4's dummy data, is constructed as:

	Control	Exp	Total
Correct	11	7	18
Incorrect	4	8	12
Total	15	15	30

These comprise the _observed_ values. The _expected_ values are computed in the following manner:

$$Correct_{control} = \frac{18 * 15}{30} = 9$$

$$Correct_{experim} = \frac{18 * 15}{30} = 9$$

$$Incorrect_{control} = \frac{12 * 15}{30} = 6$$

$$Incorrect_{experim} = \frac{12 * 15}{30} = 6$$

53

The chi-squared value is computed according to this formula (25:581):

$$X^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}} \qquad (11)$$

However,

> When the continuous X^2 distributions are applied to discrete distributions, a correction for continuity called Yates' correction is available. The correction is made by reducing each absolute difference between observed and expected by .5. In general, Yates' correction is applied only when the number of degrees of freedom is equal to one (24:432-433).

Since the data from Table 4 comprises a discrete chi-squared distribution, and because a two-way contigency table has one degree of freedom, the Yates' correction shall be used in the analysis (24:433). Thus, the formula used to compute the chi-squared statistic is (24:433):

$$X^2 = \sum \frac{(|\text{observed} - \text{expected}| - .5)^2}{\text{expected}} \qquad (12)$$

Here, $X^2 = .250 + .250 + .375 + .375 = 1.25$. With one degree of freedom, one could reject the null hypothesis and accept the null hypothesis (the chart is misleading) at a significance level of .264. Given this large p-value, it may be better to not reject the null hypothesis.

Survey Graph Analysis

Investigative question number 3 examined if program managers/cost analysts were creating graphs which violated the criteria for creating high-integrity graphics. Several dozen graphs from throughout AFSC were returned. The analysis effort shall consist of comparing these graphs to the criteria in Tables 1 and 2 and noting any differences. All findings are reported in Chapter IV.

Summary

Commonly-used graphs of CPR data were identified and cross-referenced against the criteria for high-integrity graphs. When one of the criterion is violated, the potential exists for a distorted or misleading graph. By creating high-integrity and misleading graphs of the identical data, an experiment can be conducted to see if the misleading graphs can deceive the reader. A t-test on the differences of two means was the statistic used in the experimental analysis. In addition, two-way contigency tables may be used to examine, chart by chart, how effective each misleading graph was in fooling the reader. Chapter IV, Analysis and Findings, contains the results of the experiment.

IV. Analysis and Findings

Experimental Results

The experimental hypotheses were:

Ho: $(O_1 - O_2) - (O_3 - O_4) = 0$, the graphs were not misleading.

Ha: $(O_1 - O_2) - (O_3 - O_4) > 0$, the graphs were misleading.

Utilizing the format of table 4 (Chapter III), statistical summaries of both experimental and control groups were generated, as shown in Appendix D. Using the statistical analysis method described in Chapter III, a t-score of 6.068 was realized from the data. With 52 degrees of freedom, the null hypothesis may be rejected at an alpha of less than .0001.

Referring to table 5, the charts having the most misleading effect may be identified through the use of two-way contigency tables. With the exception of the rotated bar chart, all charts had a significance (p-value) of less than .0712, with several much lower. These levels indicate that these charts, when not constructed in accordance with the high-integrity criteria, can mislead the readers.

Analysis

Since the significance level is low, the experimental evidence indicates the alternative hypothesis (Ha), which states the graphs violating the high-integrity criteria proved misleading, should be accepted. Thus, it has been

Table 5.   Posttest Chart Two-Way Contigency Table Results

Chart	X^2 Value	p-value
1	.05	.6133
2	16.58	.0000
3	18.22	.0000
4	3.25	.0712
5	12.58	.0004
6	6.68	.0098

Notes

Chart 1 (rotated bar chart) attempted to mislead the reader by reversing the x-axis scale.

Chart 2 (area graph) attempted to mislead the reader by placing the most irregular strata at the bottom.

Chart 3 (bar chart) attempted to mislead the reader by using volumes instead of linear magnitudes.

Chart 4 (line chart) attempted to mislead the reader by reversing the vertical scale.

Chart 5 (line chart) attempted to mislead the reader by reversing the
x-axis time scale.

Chart 6 (bull's-eye chart) attempted to mislead the reader by mislabeling quadrants.

demonstrated that Air Force decision makers can be misled by low-integrity graphics. This is the overall summary of the experiment. However, as mentioned above, not all charts had the same misleading effect on the experimental group students.

Referring back to Table 5, the results indicate the students were misled the most by charts 2, 3, and 5, and not misled by chart 1. The large disparity between charts 1 and 5 is perplexing, as both misleading versions were constructed in the same manner (the horizontal scales were reversed). One may expect the results of these two charts to be similar. A possible explanation for this is the degree of learning which took place during the experiment. More is said about this in the Experimental Issues section of this chapter.

Survey Graph Analyses

Graphs were received from the Electronics Systems Division, Aeronautical Systems Division, Ballistics Systems Division, and Munitions Systems Division. Human Systems Division explained their programs do not meet the threshold for CPR use, and therefore were unable to comply with the request for graphs. Space Systems Division did not respond.

Various kinds of graphs of CPR data were received. Line graphs portraying Cumulative Contract Status and cost/schedule performance indices, and bar graphs depicting cost and schedule variances were the most common. None of

the more difficult charts to read like "Bull's-eye" or area graphs were received. To what extent they are used cannot be determined, nor was it an area researched in this effort. It appears, from the limited amount of charts received, that the simpler graphs to create and comprehend (i.e., line and bar graphs) are the most widely used.

Most graphs were constructed in a high-integrity manner. Several charts were distorted by not using a zero baseline. Appendix E contains selected graphs that violated at least one of the criteria, along with an explanation of how each could be constructed in accordance with the high-integrity criteria.

Nearly all charts received were computer generated. Some were done by spreadsheet-type software systems, while others appear to have been produced by automated analysis programs like the Space Systems Division "Performance Analyzer", the Air Force Cost Center "CPR-EZ", and the Air Force Cost Center "Cost Analysis System". Only one graph received contained hand-plotted data. It is included in Appendix E as an example of what not to do, because it does not accurately reflect the underlying data. Overall, the assumption that many charts are being constructed using computer graphics software, and that some of these graphs may violate high-integrity criteria, was supported by the evidence.

Experimental Issues

As mentioned previously, the experimental design was:

$$R \quad O_1 \quad X \quad O_2$$

$$R \quad O_3 \quad X \quad O_4$$

This design calls for the application of a pretest, followed by the posttest. However, in both trials, the intended pretest was mixed in with the posttest, which effectively changed the design to Posttest-Only:

$$R \quad X \quad O_1$$

$$R \quad \quad O_2$$

This error was noted after the second trial. However, the same method was used for trial 2 to yield a large amount of data when pooled with the data from trial 1. A statistical analysis was performed on the data as a Posttest-Only design. The analysis yielded a t-score of -8.1841 with 55 degrees of freedom. This result still shows a significant difference (at a significance level of less than .0001) in the number of correct responses, indicating the misleading charts impaired the students' ability to comprehend the data being portrayed.

As mentioned in Chapter III, the only factor that is a major weakness of the Pretest-Posttest Control Group Design is the reactive or interaction effect of testing, in which a pretest might increase or decrease the respondent's sensitivity or responsiveness to the experimental variable. After each experiment was administered, the students were given a chance to provide oral and written feedback.

60

Several students said they learned to look more closely at the graphs, especially each axis, as the experiment progressed. Had this reactive effect not taken place, the difference between the two groups could only be greater.

V. Conclusion

Graphs are a very quick and easy way of portraying
tabular data. In the Air Force, CPR data can easily be
graphed using spreadsheet software and desktop computers.
The ease with which this may be done invites the temptation
to make the data "look better" than it actually is. Even
when there is no _deliberate_ attempt to misrepresent the
data, powerful graphics software enables the user to
unwittingly distort the messge. There are many ways to do
this. Most methods take advantage of the reader's
assumption that "everything is where it should be"--proper
labeling, each axis normally constructed, and the overall
graph being simple to add to clarity, among others. The use
of misleading graphics may have very ominous results.

Summary of Results

A review of the literature confirmed the existence of
criteria for high-integrity graphs. In addition, there are
many "style guides" which, if followed, may help make graphs
portray information more clearly. While there are no laws
stipulating the use of these guides, they serve to aid the
makers of graphs to convey information more clearly.

Sixty-three students participated in the experiments to
determine if decision makers could be affected by misleading
graphics. The results showed that decision-makers _can_ be
misled with statistical significance. However, this in no
way indicates that they _are_ being misled in their jobs.

62

Also, while program managers and analysts are creating some graphs which distort data (shown in Appendix E), there is no evidence of intent to mislead the users of the graphs.

Not all charts had the same misleading effect. Area graphs, bar charts (using volumes instead of linear magnitudes), line charts, and "bulls-eye" charts, when constructed in ways which violated the high-integrity criteria, proved misleading at a significance level of less than .0712. The students were not misled by the rotated bar chart, as the significance level of .8290 clearly indicates (see Table 5, page 57).

Recommendations for Future Research

While defense contractors are not always required to submit graphs with CPR data, they are often included in program review books provided to government personnel. The graphs not only portray CPR data, but other information such as hours worked, scrap rates, etc. Are these graphs constructed in accordance with the high-integrity criteria? Is there some intent to mislead the users? Because program offices are often viewed as advocates, the possibility of deliberate intent to deceive is not remote. This may be one area for further research.

Another area to research would be different graphs used in CPR analysis. This experiment used only six of the many graphs widely used by cost analysts and program managers. Many others are in use. Also, it would be much easier to
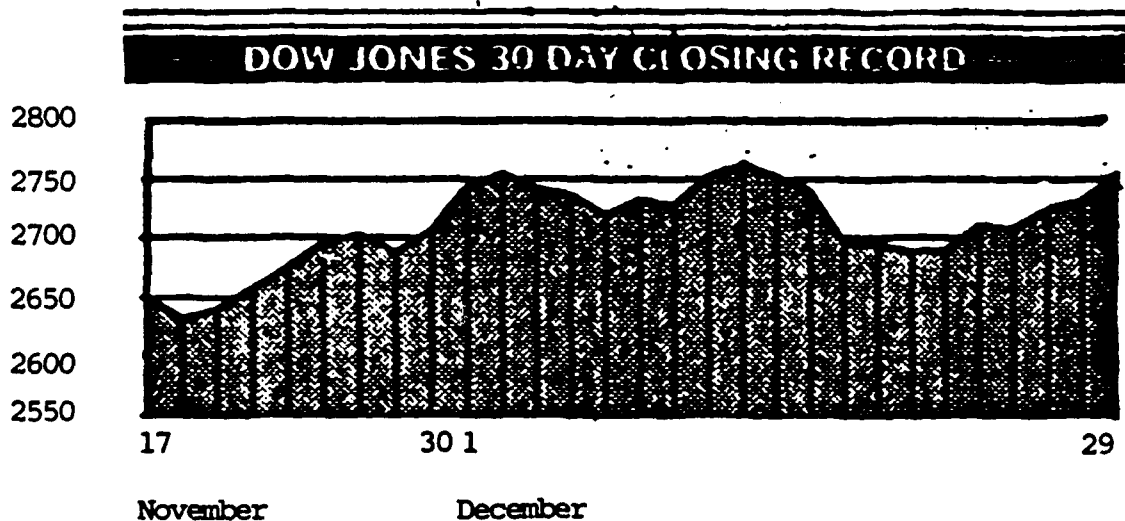
computerize the entire experiment. Each student could input his/her response at a terminal. The cost of reproducing all charts could be avoided, and a program to statistically analyze the results could be built in. This would save many hours, greatly simplify the experiment, and improve the experiment's external validity, as graphs can now be viewed directly from the computer screen.

$$\frac{2753 - 2650}{2650} * 100\% = 3.89\% \text{ (size of effect in data)}$$

$$\frac{2.3 \text{ cm} - 1.15 \text{ cm}}{1.15 \text{ cm}} * 100\% = 100\% \text{ (size of effect in graphic)}$$

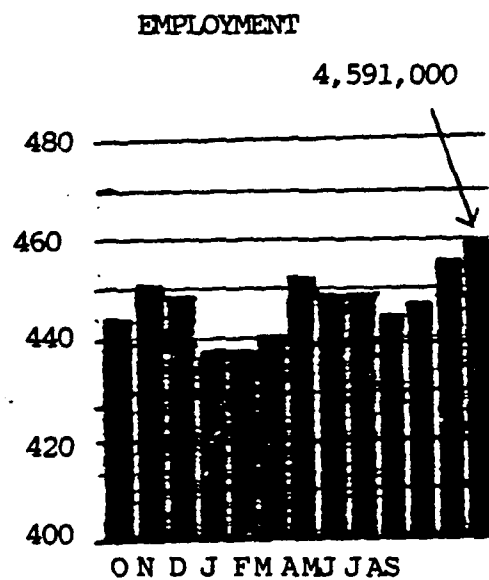$$\text{Lie Factor} = \frac{100\%}{3.89\%} = \underline{25.7}$$

DOW JONES 30 DAY CLOSING RECORD

2800			
2750			
2700			
2650			
2600			
2550			
17	30 1		29

November   December

Reprinted from Dayton Daily News, 30 Dec 89

65

$$\frac{4,591,000 - 4,455,000}{4,455,000} * 100\ \% = 3.05\% \text{ (size of effect in data)}$$

$$\frac{3.28\ cm - 2.5\ cm}{2.5\ cm} * 100\ \% = 31.2\ \% \text{ (size of effect in graphic)}$$

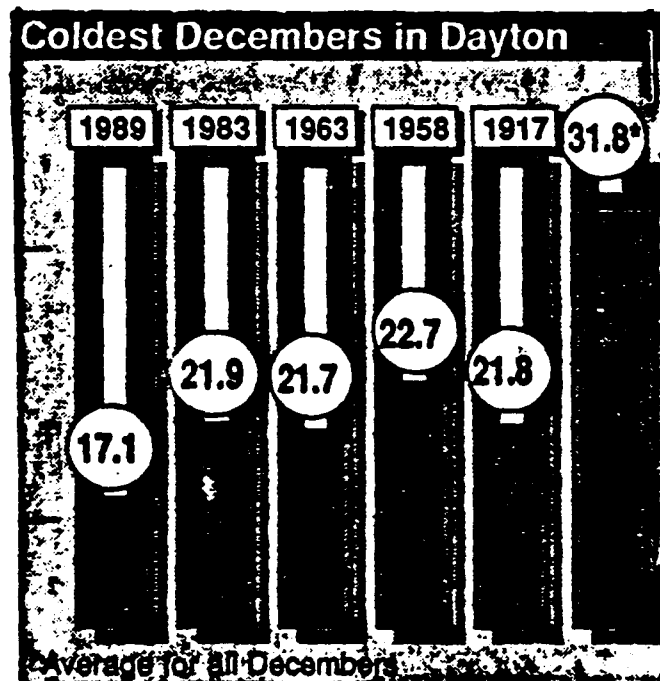$$\text{Lie Factor} = \frac{31.2}{3.05} = \underline{10.23}$$

EMPLOYMENT

4,591,000

Reprinted from Dayton Daily News, 26 Dec 89

$$\frac{31,8 - 17.1}{17.1} * 100 \text{ \%} = 86 \text{ \%} \text{ (size of effect in data)}$$

$$\frac{7 \text{ cm} - 2.1 \text{ cm}}{2.1 \text{ cm}} * 100 \text{ \%} = 233.33 \text{ \%} \text{ (size of effect in graphic)}$$

$$\text{Lie Factor} = \frac{233.33 \text{ \%}}{86 \text{ \%}} = \underline{2.71}$$

Coldest Decembers in Dayton

1989   1983   1963   1958   1917   31.8*

17.1   21.9   21.7   22.7   21.8

*Average for all Decembers

Reprinted from Dayton Daily News, 1 Jan 90

DEPARTMENT OF THE AIR FORCE
HEADQUARTERS AIR FORCE SYSTEMS COMMAND
ANDREWS AIR FORCE BASE DC 20334-5000

28 NOV 1989

REPLY TO
ATTN OF   AC

SUBJECT.   Thesis Research Material Request

TO   ASD/AC   BSD/AC   ESD/AC   HSD/AC   MSD/AC   SSD/AC

1. Research is currently underway at the Air Force Institute of Technology
to analyze Cost/Schedule Control Systems Criteria (C/SCSC) Cost Performance
Report (CPR) graphical analysis techniques. This effort is specifically
directed at the effect misleading graphics may have on CPR analysis. The
relationship between the graphical display of data, the information
presented to managers, and the decisions resulting from CPR analysis
techniques will be considered. To gain an understanding of the way CPR
data is portrayed within the product divisions, it is important to gather
examples of both CPR data and the corresponding graphics.

2. The goal of this research is to identify ways in which managers use CPR
graphs for comparative purposes. By understanding the ways these graphs
are used, we then could understand their influence in the program office
and the extent of their use for program decisions. To this end, many
samples of graphs and data from various programs are needed.

3. Request your office forward examples of CPR data and related graphics
from your programs as prepared by program manager/program control
personnel. All we wish is the CPR data and "in-house" graphics
corresponding to the data. These materials should be mailed to the
following address:

     AFIT/LSG
     Attn: Capt Larkin
     Wright-Patterson AFB OH 45433-6583

Please ensure all data are unclassified. It is requested that material be
provided by 31 December 1989.

4. Your cooperation is greatly appreciated. Please address any questions
you may have to Capt Albert Larkin at AV 785-4437.

JOHN M. NAUSEEF
Brigadier General, USAF,
DCS/Comptroller

Cover Letter

Research is underway at AFIT to determine how graphs help cost
analysts and program managers perform their jobs more effectively.
For example, a long list of data may not quickly convey information
managers and cost analysts need. However, by graphing this data, they
can quickly see trends or areas that need special attention. Today,
with microcomputers on virtually every manager's desk, it is very easy
to transform data into very useful and informative graphs. But how
much is it really helping them?

We would like you to take part in an experiment to help us answer
this question. This package contains twelve graphs which summarize
selected cost and schedule information taken from Cost Performance
Reports of the B-3 program. Above each graph is a short explanation
of how it is used by managers and cost analysts. Please form an
IMPRESSION in your mind of what information the graph is portraying.

The students should review the graph for 30 seconds, and then
turn the page. On the next page will be a conclusion based on the
graph you just looked at. Please indicate whether you agree or
disagree with the conclusion by circling the appropriate answer.
15 seconds will be provided for the students to read the conclusion
and circle their answer. A relatively short amount of time is
allotted for this, as we are trying to determine, as realistically as
possible, the impressions managers and cost analysts make from cost
and schedule type graphs. The total length of this experiment should
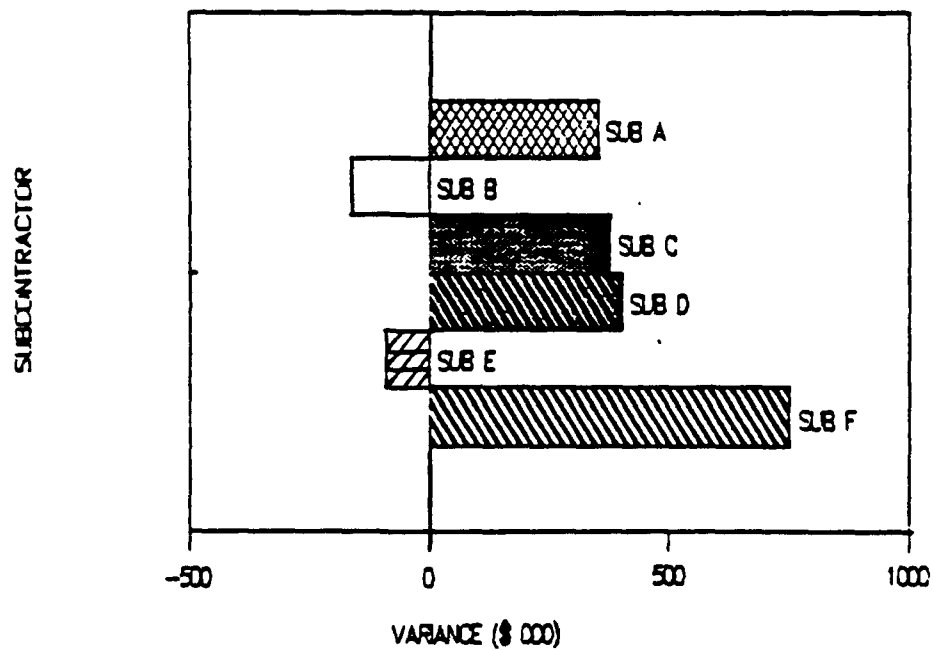not exceed 20 minutes.

Your cooperation in this effort is greatly appreciated. Thank
you.

CUMULATIVE SUBCONTRACTOR COST VARIANCES ($000)

Purpose: This graph illustrates the cumulative cost performance of
selected B-3 bomber subcontractors.

B-3 SUBCONTRACTOR

CUMULATIVE COST VARIANCES

Conclusion: Four subcontractors are ahead of cost and two are
behind cost on this project.
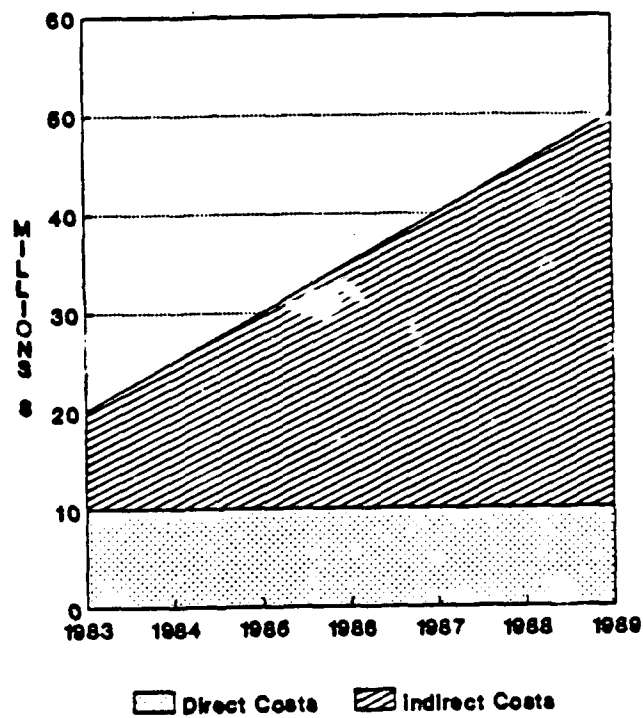
DISAGREE   AGREE

B-3 DEFENSIVE AVIONICS SUBCONTRACTOR
ANNUAL DIRECT AND INDIRECT COSTS

Purpose: This graph shows how total annual costs are divided into
direct and indirect costs from 1983 through 1989.

Annual Costs

72

Conclusion: Annual direct costs are increasing throughout the
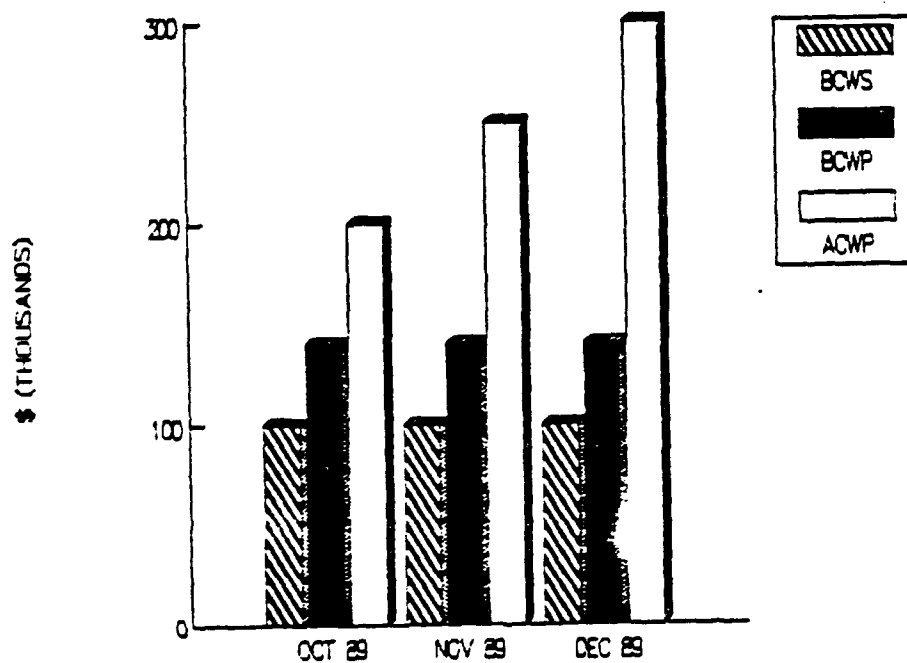period 1983 to 1989.

DISAGREE   AGREE

MONTHLY COMPARISON OF B-3 ENGINE
SUBCONTRACTOR'S COST PERFORMANCE
(4TH QUARTER CALENDAR YEAR 1989)

Purpose: This graph tracks the budgeted and actual costs of work

MONTHLY COST DATA

74

Conclusion: The monthly ACWP is constant throughout the period shown.

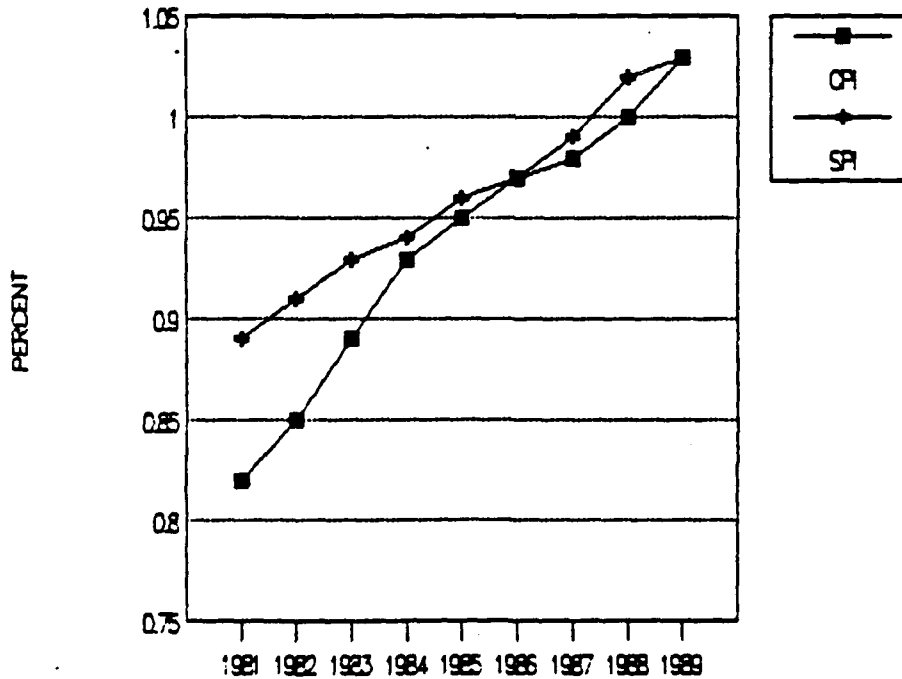   DISAGREE   AGREE

B-3 WEAPONS DELIVERY GUIDANCE SYSTEM
SUBCONTRACTOR PERFORMANCE INDEX TRENDS

Purpose: This graph compares the schedule performance index for
efficiency to the cost performance index for efficiency. The
following information may prove helpful:

SPI = cumulative BCWP ÷ cumulative BCWS

CPI = cumulative BCWP ÷ cumulative ACWP

PERFORMANCE INDEX TRENDS .

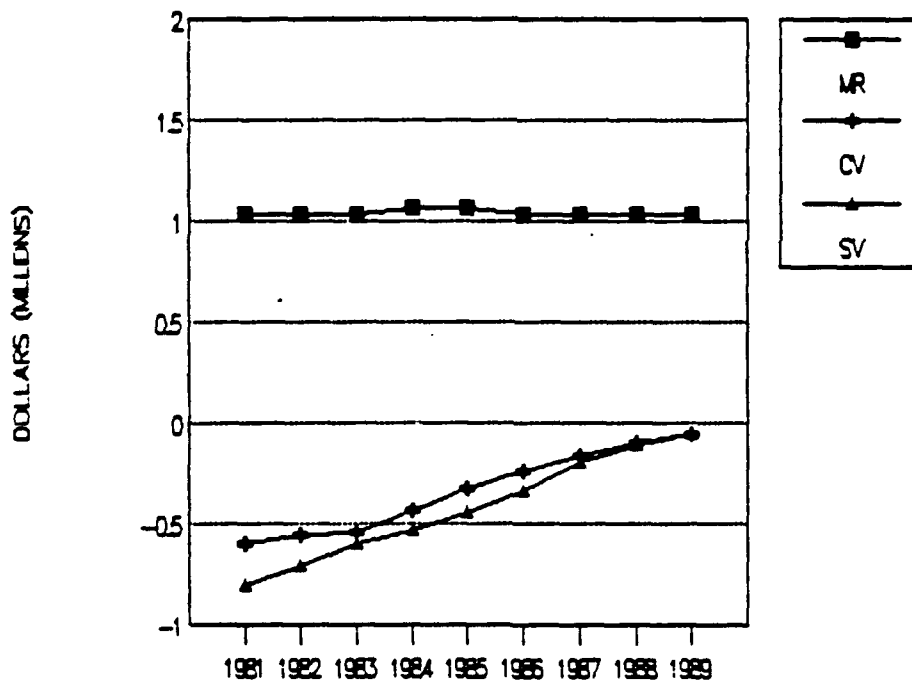Conclusion: The CPI and SPI are at their best levels in 1989.

             DISAGREE   AGREE

Purpose: This graph shows cost and schedule variance trends and the level of management reserve.

MR LEVEL AND C/S VARIANCES

Conclusion: The cost and schedule variances have steadily risen
and were at their best levels in 1989.

                    DISAGREE   AGREE
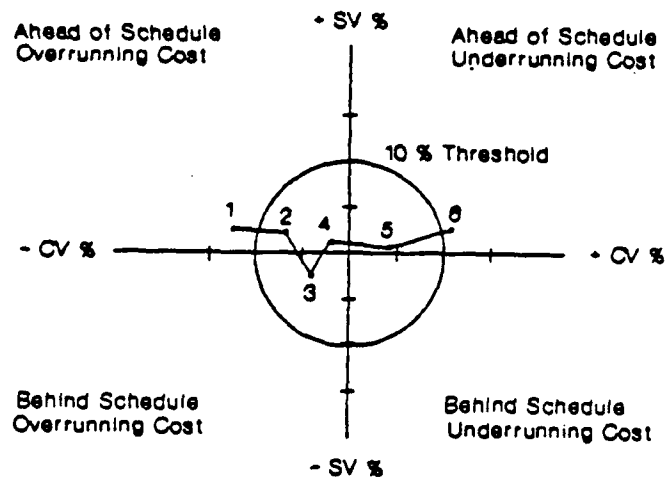
B-3 NAVIGATION SYSTEM SUBCONTRACTOR
CUMULATIVE COST AND SCHEDULE VARIANCE PERCENTAGE
(1 JULY 1989 TO 31 DEC 1989)

Purpose: This graph illustrates the cumulative cost and schedule
variance trends for a major B-3 subcontractor.

CUMULATIVE COST AND SCHEDULE VARIANCE PERCENTAGE

Conclusion: This program is behind schedule and underrunning cost.

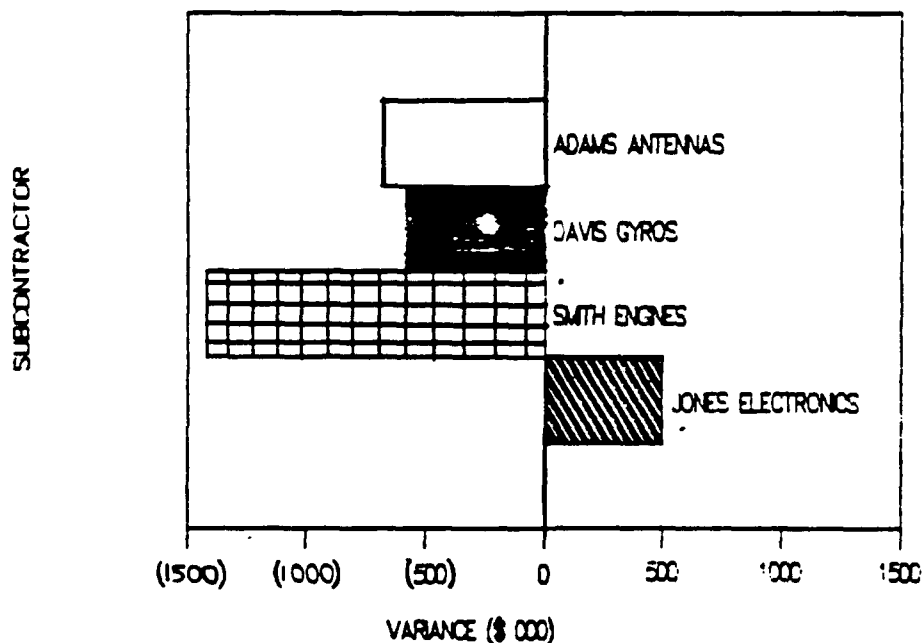DISAGREE   AGREE

CUMULATIVE SUBCONTRACTOR VARIANCES

Purpose: The purpose of this graph is to portray the cumulative
schedule performance of selected major subcontractors on the B-3
program.

B-3 SUBCONTRACTOR VARIANCES

(CUMULATIVE)

Conclusion: Of the subcontractors shown, Jones Electronics has the
largest negative schedule variance.

        DISAGREE   ·   AGREE

B-3 SPO ANNUAL SUPPORT COSTS
(1980 - 1986)

Purpose: This graph shows how total costs in the B-3 SPO are broken
out into direct and indirect costs.

Annual Direct and Indirect Costs

(SMIL)

$30
$25
$20
$15
$10
$5
$0

1980   1981   1982   1983   1984   1985   1986

Indirect Costs   Direct Costs

84

Conclusion: Indirect costs have been the same each year.

        DISAGREE   AGREE

MONTHLY COMPARISON OF B-2 LANDING GEAR
SUBCONTRACTOR'S COST PERFORMANCE
(4TH QUARTER CALENDAR YEAR 1989)

Purpose: This graph is used to track the budgeted and actual costs
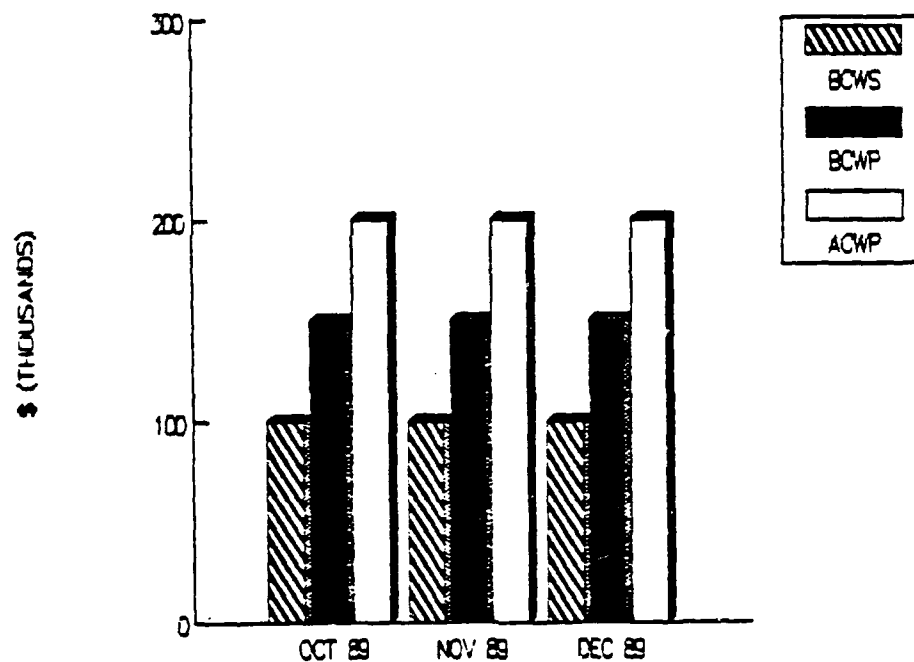of work scheduled and completed during the period.

MONTHLY COST DATA

Conclusion: The ACWP is steadily decreasing throughout the period.

DISAGREE   AGREE

B-3 MISSILE GUIDANCE SYSTEM
PERFORMANCE INDEX TRENDS

Purpose: This graph compares the schedule performance index for
efficiency to the cost performance index for efficiency. The
following information may prove helpful:

SPI = cumulative BCWP $\div$ cumulative BCWS

CPI = cumulative BCWP $\div$ cumulative ACWP

PERFORMANCE INDEX TRENDS

Conclusion: The CPI and SPI values in 1989 are better than those in
1981.

                    DISAGREE   AGREE

Purpose: This graph shows cost and schedule variance trends and the
level of management reserve in the B-3 program from 1981 to 1989.

MR COMPARISON

TO CV AND SV

Conclusion: The cost variance and schedule variance reached their
worst levels in 1989.

                  DISAGREE   AGREE

Purpose: This graph tracks both cost variance and schedule variance
trends over a six month period.

Cumulative Cost and Schedule Variance Percentage (6 Months)

Conclusion: At month 6, this program was behind schedule and overrunning cost.

DISAGREE   AGREE

CUMULATIVE SUBCONTRACTOR VARIANCES

Purpose: The purpose of this graph is to portray the cumulative
schedule performance of selected major subcontractors on the B-3
program.

B-3 SUBCONTRACTOR VARIANCES

(CUMULATIVE)

Conclusion: Of the subcontractors shown, Jones Electronics has the largest negative schedule variance.

                    DISAGREE   AGREE

B-3 SPO ANNUAL SUPPORT COSTS
(1980 - 1986)

Purpose: This graph shows how total costs in the B-3 SPO are broken
out into direct and indirect costs.

Annual Direct and Indirect Costs

Conclusion: Indirect costs have been the same each year.

       DISAGREE   AGREE

MONTHLY COMPARISON OF B-3 LANDING GEAR
SUBCONTRACTOR'S COST PERFORMANCE
(4TH QUARTER CALENDAR YEAR 1989)

Purpose: This graph is used to track the budgeted and actual costs
of work scheduled and completed during the period. Data are
non-cumulative.

MONTHLY COST DATA
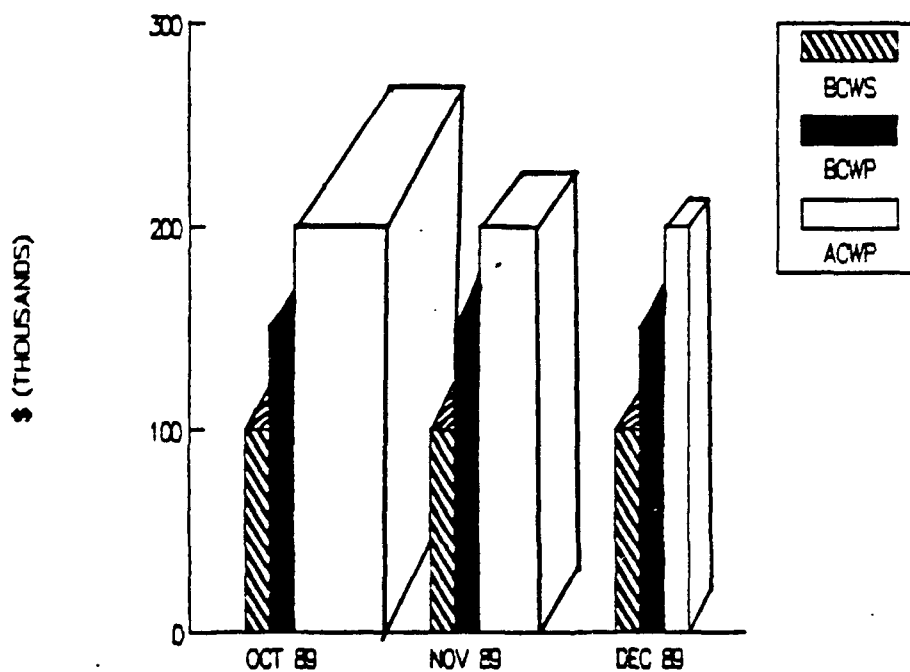
98

Conclusion: The ACWP is steadily decreasing throughout the period.

            DISAGREE   AGREE

Purpose: This graph compares the schedule performance index for
efficiency to the cost performance index for efficiency. The
following information may prove helpful:

SPI = cumulative BCWP ÷ cumulative BCWS

CPI = cumulative BCWP ÷ cumulative ACWP

PERFORMANCE INDEX TRENDS

Conclusion: The CPI and SPI values in 1989 are better than those
in 1981.

                    DISAGREE   AGREE

MANAGEMENT RESERVE USAGE COMPARED
WITH COST VARIANCE AND SCHEDULE VARIANCE

Purpose: This graph shows cost and schedule variance trends and the
level of management reserve in the B-3 program from 1981 to 1989.

MR COMPARISON

TO CV AND SV

102

Conclusion: The cost variance and schedule variance reached their
worst levels in 1989.

DISAGREE   AGREE

Purpose: This graph tracks both cost variance and schedule variance
trends over a six month period.

Cumulative Cost and Schedule Variance Percentage (6 Months)

+ SV %

Ahead of Schedule   Ahead of Schedule

Underrunning Cost   Overrunning Cost

10 % Threshold

1

2

- CV %   + CV %

6   5

3

4

Behind Schedule   Behind Schedule

Underrunning Cost   Overrunning Cost

- SV %

Conclusion: At month 6, this program was behind schedule and
overrunning cost.

                    DISAGREE   AGREE

Appendix D: Experiment Statistical Summaries

Control Group Statistical Summary

Student	P1	P2	P3	P4	P5	P6	Tot	1	2	3	4	5	6	Tot	Delta	d-sqr
	03							04									
1	1	1	1	1	0	1	5	1	1	1	1	0	1	5	0	0.0375
2	1	1	1	1	1	0	5	1	1	1	0	1	0	4	1	1.4246
3	0	1	1	1	1	1	5	1	1	1	1	1	1	6	-1	0.6504
4	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
5	1	1	1	1	0	1	5	1	1	1	1	1	1	6	-1	0.6504
6	1	0	1	1	0	1	4	1	1	1	1	1	1	6	-2	3.2633
7	1	0	1	1	1	1	5	0	1	1	1	1	1	5	0	0.0375
8	1	1	1	0	1	1	5	1	1	1	1	1	1	6	-1	0.6504
9	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
10	1	1	1	0	1	1	5	1	1	1	1	1	1	6	-1	0.6504
11	1	1	1	0	1	1	5	1	1	1	1	1	1	6	-1	0.6504
12	1	1	1	0	0	1	4	1	1	1	1	1	0	5	-1	0.6504
13	1	1	1	1	0	1	5	0	1	1	1	1	1	5	0	0.0375
14	1	1	1	0	0	1	4	1	1	1	1	1	0	5	-1	0.6504
15	0	1	0	1	1	1	4	0	1	1	1	1	1	5	-1	0.6504
16	1	1	0	1	1	1	5	1	1	1	0	1	1	5	0	0.0375
17	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
18	1	1	1	1	1	1	6	1	1	1	1	1	0	5	1	1.4246
19	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
20	1	1	1	1	1	1	6	1	1	1	1	0	1	5	1	1.4246
21	1	1	1	0	1	1	5	0	1	1	1	1	1	5	0	0.0375
22	1	0	1	0	1	1	4	0	1	1	1	1	1	5	-1	0.6504
23	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
24	1	1	1	1	1	1	6	1	1	1	0	1	1	5	1	1.4246
25	1	1	1	1	1	1	6	0	1	1	1	1	0	4	2	4.8117
26	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
27	1	1	1	1	0	1	5	0	1	1	1	1	1	5	0	0.0375
28	1	1	1	0	1	1	5	1	1	1	1	1	1	6	-1	0.6504
29	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
30	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
31	1	1	1	1	1	1	6	1	1	1	1	1	1	6	0	0.0375
Totals:	29	28	29	23	24	30	163	24	31	31	28	29	26	169	-6	20.839

Mdelta = -0.194

Experimental Group Statistical Summary

Student	01							02							Delta	d-sqr
	P1	P2	P3	P4	P5	P6	Tot	1	2	3	4	5	6	Tot		
1	0	1	1	1	1	0	4	0	1	0	1	0	0	2	2	0.22
2	1	1	0	1	1	1	5	1	1	0	0	0	0	2	3	2.157
3	1	1	0	1	1	1	5	1	0	0	1	0	0	2	3	2.157
4	1	1	1	1	1	1	6	1	1	1	0	0	0	3	3	2.157
5	1	1	1	0	1	1	5	1	0	0	0	0	1	2	3	2.157
6	1	1	1	0	1	0	4	0	1	1	1	1	0	4	0	2.345
7	1	1	1	1	0	1	5	1	0	0	1	1	1	4	1	0.282
8	1	1	1	0	1	0	4	1	0	0	0	0	1	2	2	0.22
9	0	0	0	1	1	1	3	0	0	1	1	1	0	3	0	2.345
10	1	1	1	1	0	1	5	1	1	1	1	1	1	6	-1	6.407
11	0	1	1	1	0	1	4	0	1	0	0	0	0	1	3	2.157
12	1	1	1	1	1	1	6	1	1	0	0	1	1	4	2	0.22
13	1	0	1	1	1	0	4	0	0	1	0	0	1	2	2	0.22
14	1	1	1	1	1	1	6	1	1	0	1	1	1	5	1	0.282
15	1	1	1	0	1	1	5	0	1	0	1	1	1	4	1	0.282
16	1	1	0	1	1	1	5	1	0	0	1	0	1	3	2	0.22
17	1	1	1	1	1	1	6	1	1	1	1	1	0	5	1	0.282
18	1	1	1	1	1	1	6	1	1	1	1	1	0	5	1	0.282
19	1	1	1	1	0	1	5	1	1	0	1	1	0	4	1	0.282
20	0	1	1	1	1	1	5	1	0	1	1	1	0	4	1	0.282
21	1	1	1	1	1	1	6	1	0	1	0	0	1	3	3	2.157
22	1	0	1	0	1	1	4	0	1	1	1	0	0	3	1	0.282
23	0	1	1	1	0	1	4	1	0	1	1	0	1	4	0	2.345
24	1	1	1	1	1	1	6	1	0	0	1	0	0	2	4	6.095
25	1	1	1	1	1	1	6	1	0	1	0	0	1	3	3	2.157
26	1	1	1	1	0	1	5	0	0	1	0	0	0	1	4	6.095
27	1	1	1	1	1	1	6	1	1	0	1	1	1	5	1	0.282
28	1	1	1	0	0	0	3	0	1	1	1	1	0	4	-1	6.407
29	1	1	1	1	1	1	6	1	0	1	1	0	1	4	2	0.22
30	1	1	1	1	1	1	6	1	1	0	1	1	1	5	1	0.282
31	1	1	1	0	1	1	5	1	0	0	1	1	1	4	1	0.282
32	1	0	1	1	1	0	4	1	1	1	1	1	0	5	-1	6.407
Totals:	27	28	28	25	25	26	159	0 23	17	16	22	16	16	110	49	57.97

Mdelta = 1.531

Appendix E: Survey Graph Analysis

This appendix contains samples of the graphs submitted from throughout AFSC in support of this thesis.

Chart E-1 is a graph of several different cost and schedule drivers constructed as a bar chart. If the chart was made correctly, each bar's length would be proportional to the size of the data. However, this graph contains bars which are longer than or not as large as they should be. Arbitrarily choosing the second column (System Eng/PM), which has a length of 3.6 centimeters (cm), as a baseline, and solving for x,

$$\frac{3.6}{\$145,000} = \frac{X}{\$100,000}$$

$$X = 2.48 \text{ cm}$$

Thus, for every $100,000 of data, each bar should be 2.48 cm long. The following table illustrates the distortion in each column:

Column	Data Size	Bar Length	Theoretical Bar Length	% Distortion
1	$143,000	3.15 cm	3.55 cm	12.70
2	145,000	3.60	3.60	0
3	347,000	9.50	8.61	9.37
4	46,000	.90	1.14	26.76
5	215,000	5.50	5.33	3.09

Had a vertical scale been used, the distortion in this chart could have been avoided.

COST/SCHEDULE VARIANCE DRIVERS

COST

SCHEDULE

(143,000) GUIDANCE CONTROL

(145,000) SYSTEM ENG/PM

(347,000) SOFTWARE

(46,000) FUNCTIONAL DESIGN

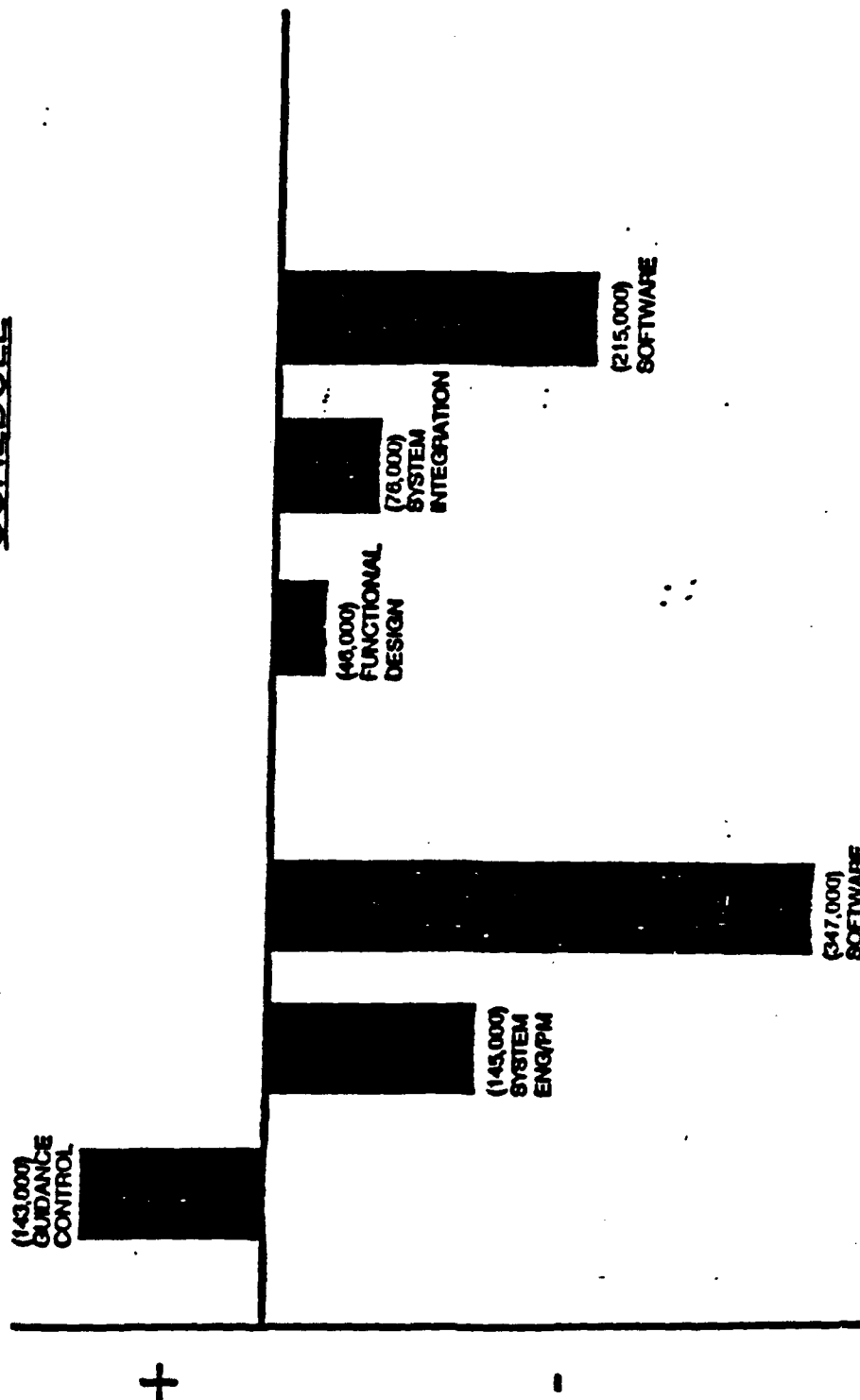(76,000) SYSTEM INTEGRATION

(215,000) SOFTWARE

+

'Chart E-1

Chart E-2 violates Criterion 10 from Table 2 (Patterned shadings should be of good contrast). The shadings here do not contrast well, and make the chart much more difficult to comprehend. A better way may have been to make one bar solid, the next lightly crosshatched, the next empty, the next heavily crosshatched, and the last a light diagonal shading. Thus, the contrast would have been much better. An even better chart would use contrasting colors.

Chart E-3 contains a distorted image of the data caused by a non-existent zero baseline. Using the Tufte Lie Factor described in Chapter 2, th e amount of distortion, utilizing the ACWP data, is:

$$\frac{2.722 - 1.40}{1.40} * 100 = 94.43 \text{ \% (size of effect in data)}$$

$$\frac{11 \text{ cm} - 1.4 \text{ cm}}{1.4 \text{ cm}} * 100 = 685.71 \text{ \% (size of efect in graphic)}$$

$$LF = \frac{685.71}{94.43} = 7.26$$

Another problem with this graph is the wide area of each entity graphed, which makes it difficult to tell exactly where the data points are.

Charts E-4 and E-5 are like E-3: both are distorted and the bands are too wide. The lie factor computation for D-4, using the cost band, is as follows:
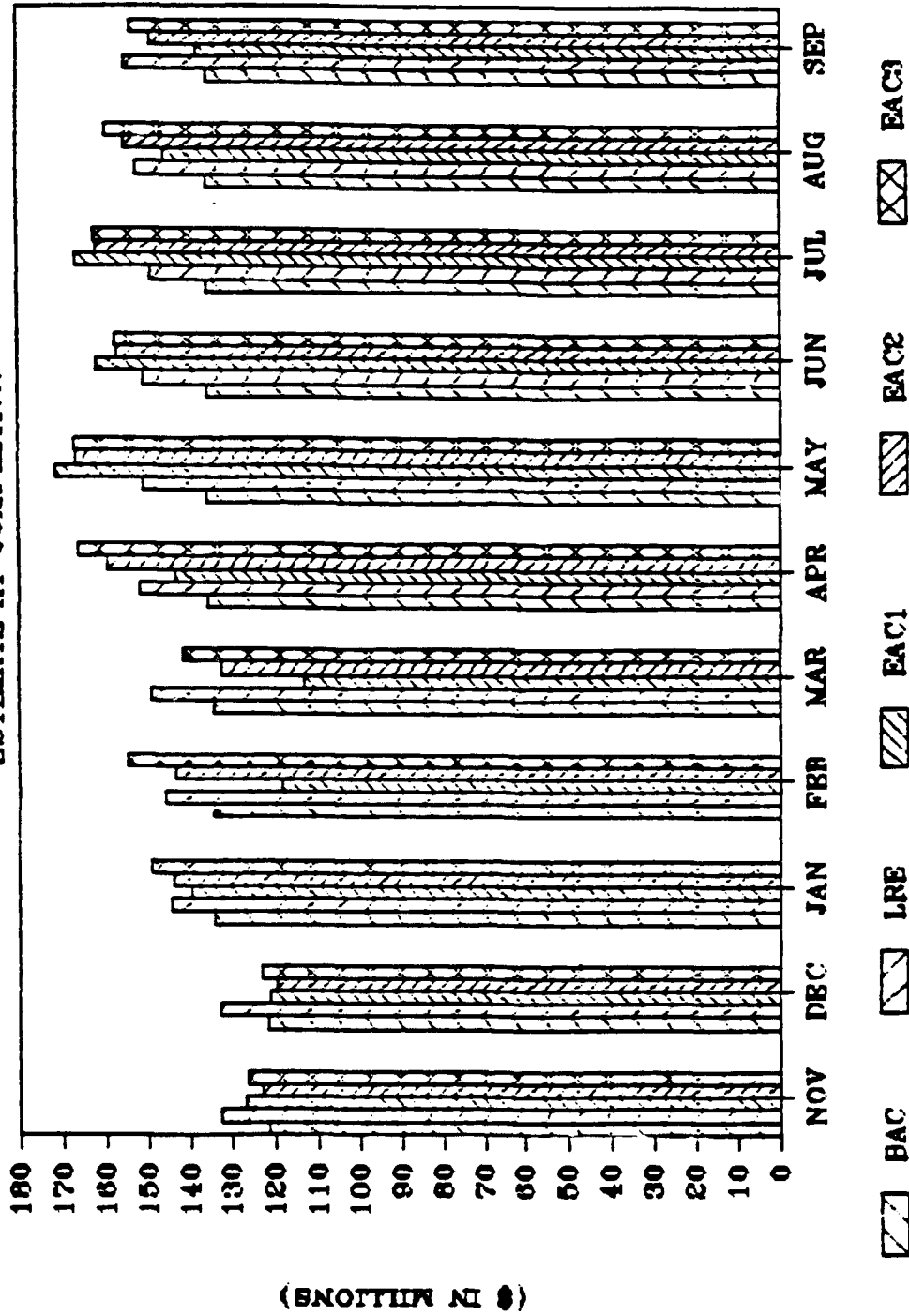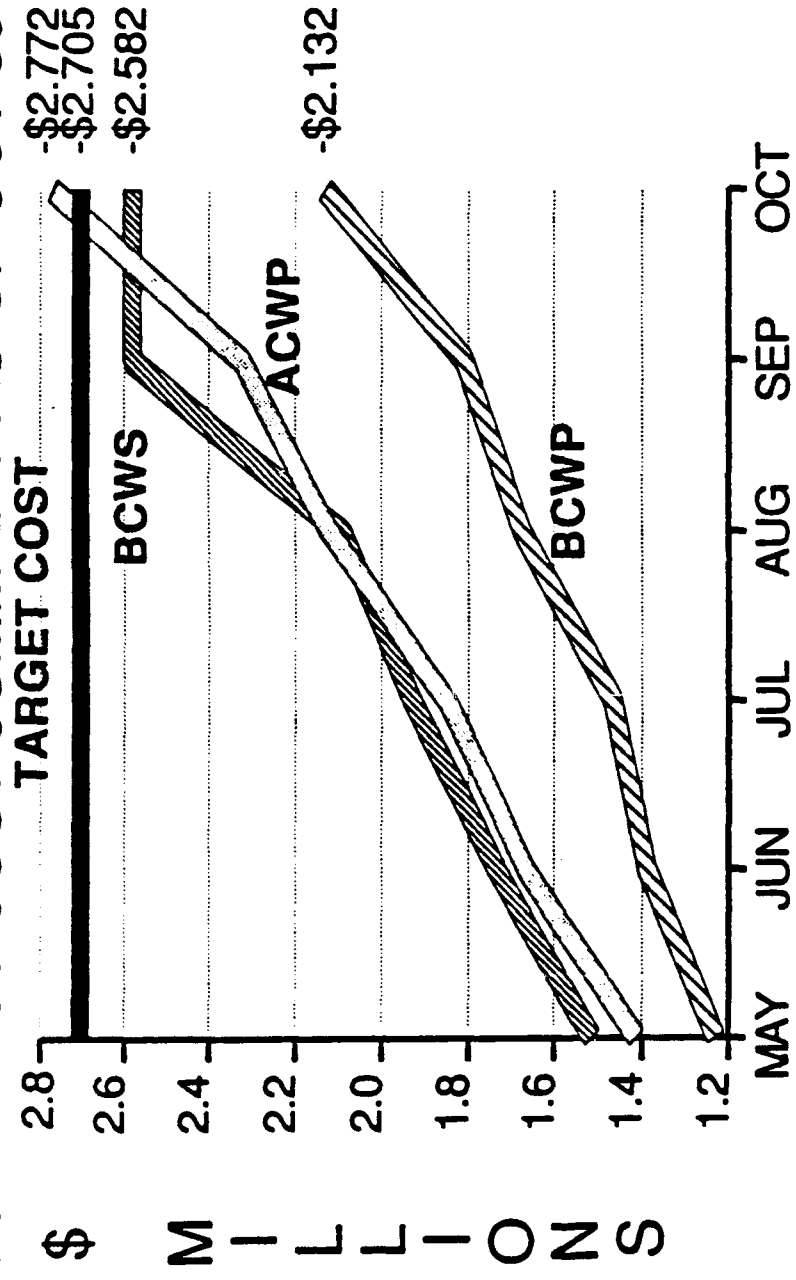
ATARS

ESTIMATE AT COMPLETION

Chart E-2

(\$ IN MILLIONS)

BAC LRE EAC1 EAC2 EAC3

NOV DEC JAN FEB MAR APR MAY JUN JUL AUG SEP

111

CONTRACT COST SUMMARY AS OF OCT 89

CPFF CONTRACT, PHASE II 82.6% COMPLETE

Chart E-3

112

COST AND SCHEDULE VARIANCE
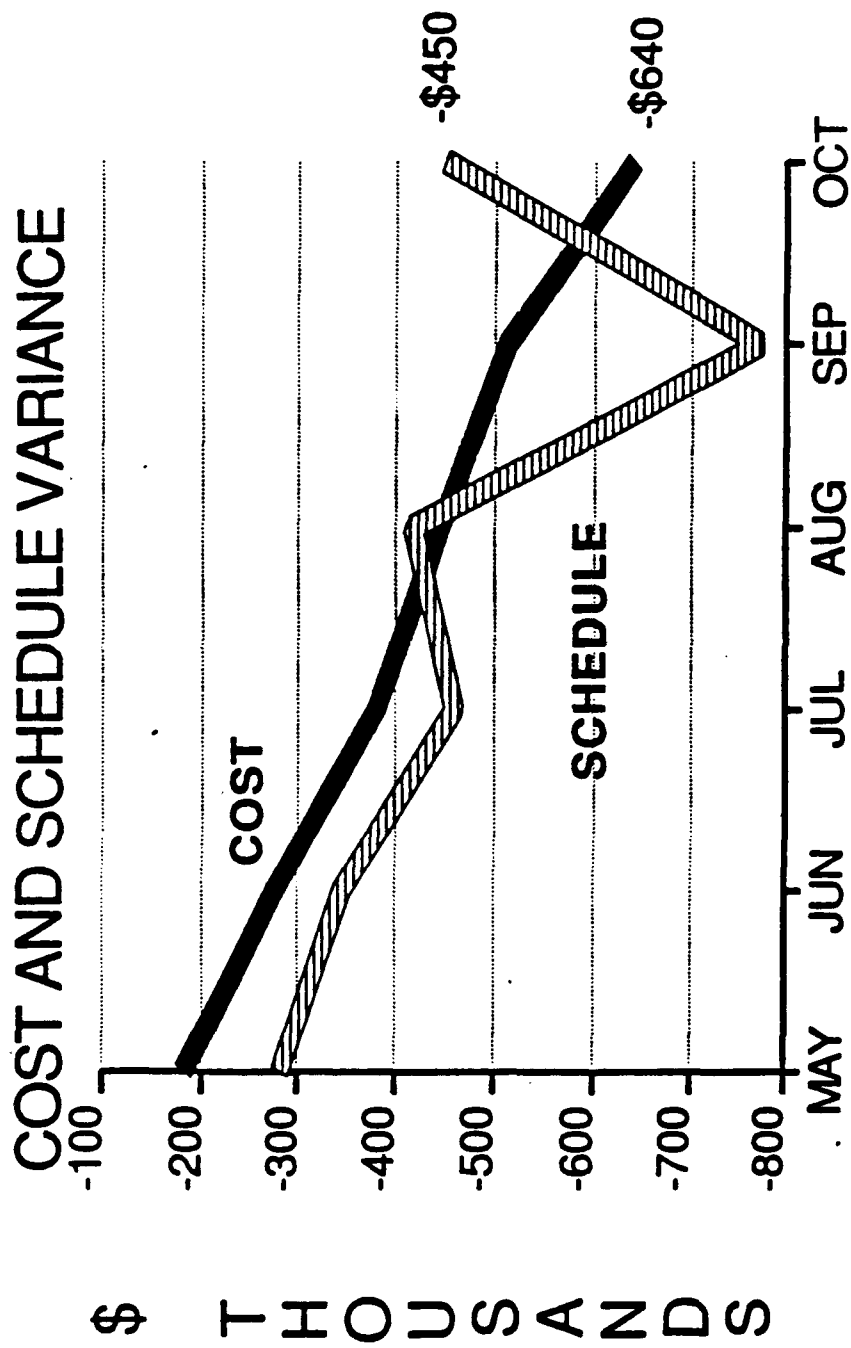
Chart E-4

PERFORMANCE INDICES

CPI

SPI

82.6

76.9

MAY JUN JUL AUG SEP OCT

88
86
84
82
80
78
76
74
72
70

Chart E-5

114
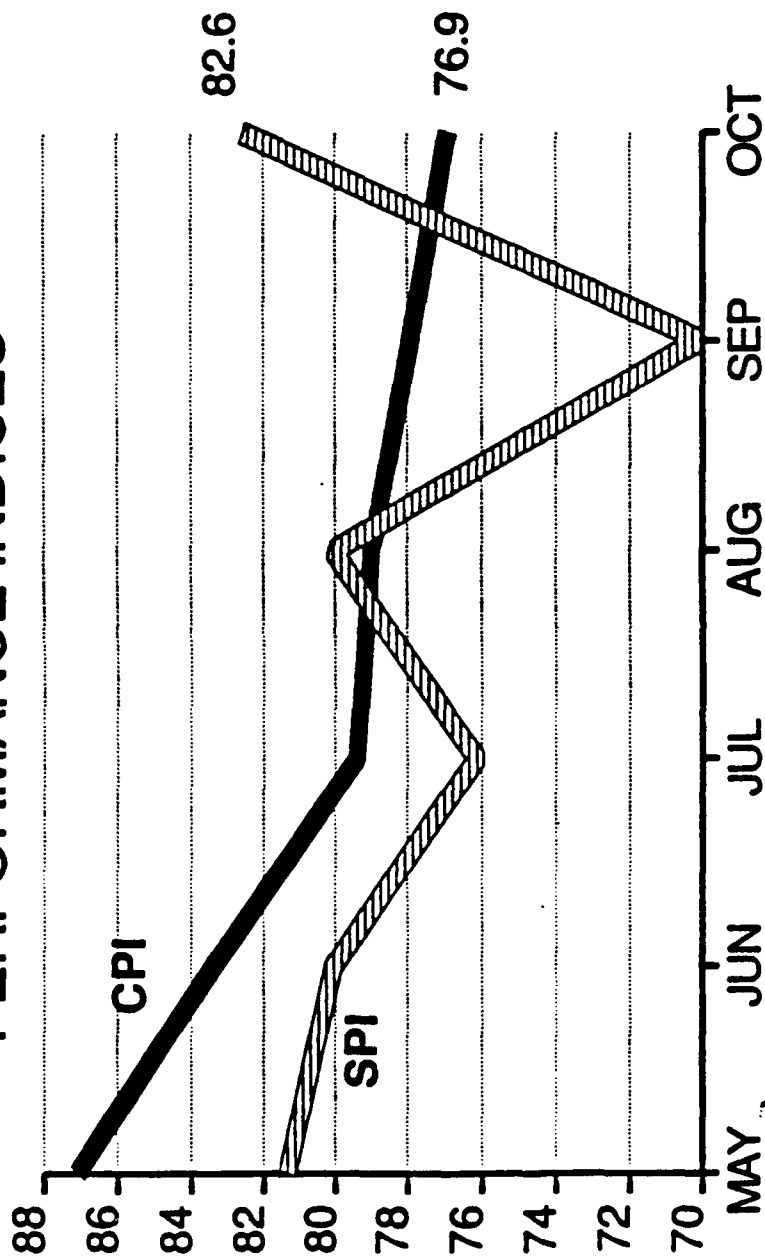
$$\frac{640 - 190}{190} * 100 = 236.84 \text{ \% (size of effect in data)}$$

$$\frac{8.55 \text{ cm} - 1.45 \text{ cm}}{1.45 \text{ cm}} * 100 = 489.65 \text{ (size of effect in graphic)}$$

$$LF = \frac{489.65}{236.84} = 2.07$$

Using the CPI band, chart E-5's lie factor is:

$$\frac{87 - 76.9}{76.9} * 100 = 13.13 \text{ (size of effect in data)}$$

$$\frac{10.1 \text{ cm} - 4.1 \text{ cm}}{4.1 \text{ cm}} * 100 = 146.34 \text{ (size of effect in graphic)}$$

$$LF = \frac{146.34}{13.13} = 11.15$$

Chart E-6 was the only graph received that appears to have data plotted by hand. The data is not distorted, but the bands are too wide and their thicknesses are inconsistent. The reader cannot tell where the data points are, and is left wondering if the relative thickness of the bands conveys additional information.

Lastly, chart E-7 suffers from a labeling problem. The legend contains almost identical symbology for both cost variance and schedule variance, leaving the reader to ponder which set of data belongs to which variance. Upon very close observation, it appears the dashed-line data reflects the schedule variance. This is true according to the table,

UNCLASSIFIED

CONTRACT COST/SCHEDULE VARIANCE GRAPH ($M)

ESTAT COMP
CONTR   SPO

FAVORABLE

UNFAVORABLE

CV%

SV%

FY QUARTER	NOV 1988	DEC 1988	JAN 1989	FEB 1989	MAR 1989	APR 1989	MAY 1989		CNTR COMPLETE	SPO COMPLETE
(BCWS) SCHED ($)	.76	1.43	2.09	2.95	3.91	4.76	5.61			
(BCWP) EARNED ($)	.72	1.28	1.87	2.59	3.66	4.40	5.22			
(ACWP) ACTUAL ($)	.73	1.38	2.00	2.72	3.72	4.58	5.55			
COST VARIANCE ($)	-.01	-.095	-.130	-.125	-.059	-.181	-.328		(1.46)	(.487)
SCHED VARIANCE ($)	-.035	-.146	-.217	-.358	-.254	-.360	-.384			

AS OF: 11 JUL 89

UNCLASSIFIED

Chart E-6

116

CONTRACT COST/SCHEDULE VARIANCE GRAPH
SRAM II/B-1B - ROCKWELL

%

12.0
10.0
8.0
6.0
4.0
2.0
0.0
-2.0
-4.0

%-COST-VARIANCE   %-SCHED-VARIANCE

	1/89	2/89	3/89	4/89	CNTR	SPO
% COST VARIANCE	6.8	9.0	8.4	10.3		
% SCHED VARIANCE	-3.3	-2.0	-0.9	1.3		
$M COST VARIANCE	0.5	0.8	1.7	2.9	0.0	0.0
$M SCHED VARIANCE	-0.3	-0.2	-0.2	0.4		
$M (BCWS)SCHED	7.8	9.4	20.7	28.1	134.6	134.6
$M (BCWP)EARNED	7.6	9.2	20.5	28.5	149.8	149.8
$M (ACWP)ACTUAL	7.1	8.4	18.8	25.6		
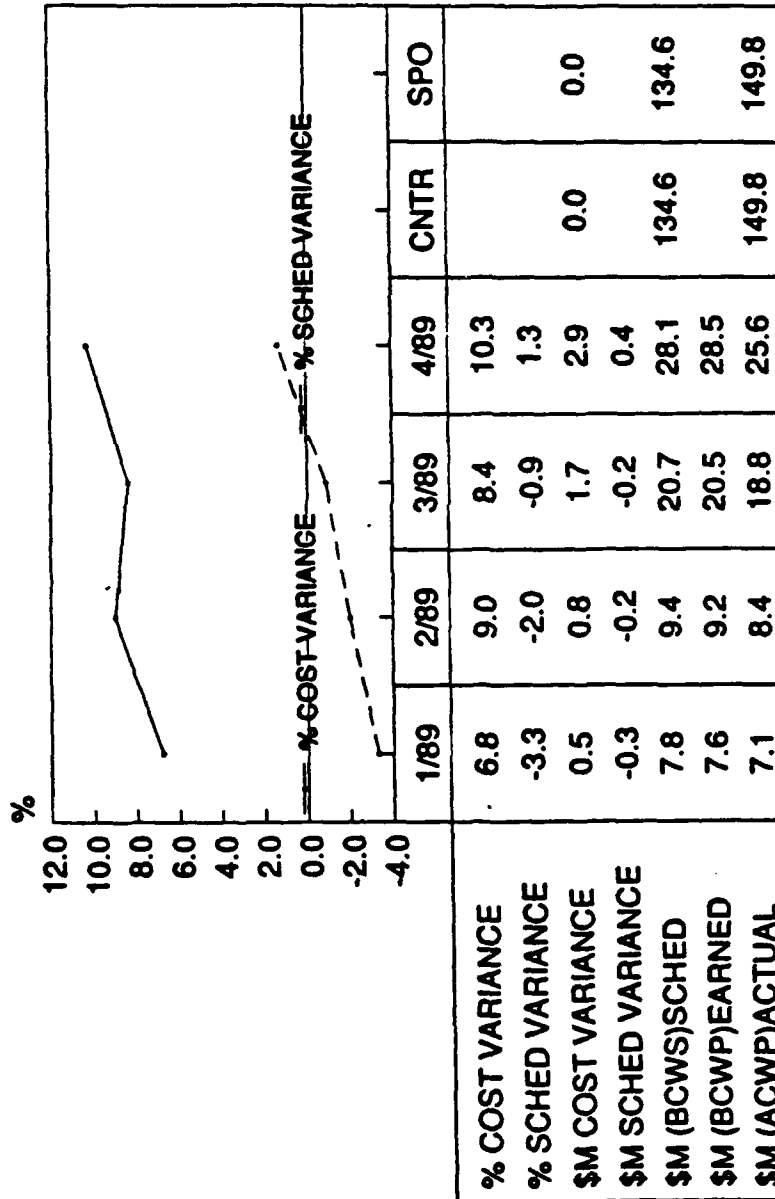
FY QUARTER

AS OF: SEP 89

Chart E-7

but tables are not always included. The real problem here is the placement of the legend. It should have been placed away from the actual data where there was no chance of the data overlaying it and making the legend hard to read.

Bibliography

1. Mintzberg, H. <u>The Nature of Managerial Work</u>, New York: Harper and Row (1973).

2. Morrocco, John D. "Cheney Proposes Stretchouts, Cuts in B-2, C-17 Programs," <u>Aviation Week and Space Technology, 18</u>: 19 (30 April 1990).

3. Fleming, Quentin W. <u>Cost/Schedule Control Systems Criteria--The Management Guide to C/SCSC</u>, Chicago: Probus Publishing Company, 1988.

4. Slemaker, Chuck M. <u>The Principles and Practice of Cost/Schedule Control Systems</u>, Princeton NJ: Petrocelli Books, 1985.

5. Department of the Air Force. <u>Cost/Schedule Control Systems Criteria Joint Implementation Guide</u>, AFSCP 173.5. Washington: HQ USAF, 1 October 1987.

6. -----. <u>Guide to Analysis ofContractor Cost Data</u>, AFSCP 173-4. Washington: HQ AFSC, 1 September 1989.

7. Spear, Mary Eleanor. <u>Practical Charting Techniques</u>, New York: McGraw Hill Book Company, 1969.

8. Tufte, Edward R. <u>The Visual Display of Quantitative Information</u>, Cheshire CT: Graphics Press, 1983.

9. Taylor, Barbara G. and Lane K. Anderson. "Misleading Graphics: Guidelines for the Accountant," <u>Journal of Accountancy 12</u>: 126-135 (October 1986).

10. Cox, D. R. "Some Remarks on the Role in Statistics of Graphical Methods," <u>Applied Statistics 27</u>: 4-9 (1978).

11. Schmid, Calvin F. and Stanton E. Schmid. <u>Handbook of Graphic Presentation</u>. New York: Ronald Press, 1954.

12. Joint Committee on Standards for Graphic Representation. "Preliminary Report Published for the Purpose of Inviting Suggestions for the Benefit of the Committee." <u>Journal of the American Statistical Association, 112</u>: 90-97 (December 1915).

13. MacGregor, A. J. <u>Graphics Simplified</u>, Toronto: University of Toronto Press, 1979.

14. Steinbart, Paul John. "The Auditor's Responsibility for the Accuracy of Graphs in Annual Reports: Some Evidence of the Need for Additional Guidance," Accounting Horizons 18: 60-70 (October 1986).

15. Johnson, Johnny R. and others. "Pictures that Lie: The Abuse of Graphs in Annual Reports," Management Accounting 10: 50-56 (October 1980).

16. Auger, B. Y. How to Run Better Business Meetings, Minnesota Mining and Manufacturing Company, 1979.

17. Rogers, Anna C. Graphic Charts Handbook, Washington: Public Affairs Press, 1961.

18. American National Standards Institute. Time Series Charts (ANSI Y15.2M), New York: The American Society of Mechanical Engineers, 1979.

19. Lefferts, Robert. Elements of Graphics: How to Prepare Charts and Graphs for Effective Reports, New York: Harper and Row, 1981.

20. Cleveland, William S. The Elements of Data Graphing, Monterey CA: Wadsworth Advanced Books and Software, 1985.

21. Feinberg, Stephen E. "Graphical Methods in Statistics," The American Statistician, 33: 165-178 (November 1979).

22. Campbell, Donald T. and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research, Chicago: Rand McNally College Publishing Company, 1966.

23. Bethea, Robert M. and others. Statistical Methods for Engineers and Scientists, New York: Marcel Dekker, 1986.

24. Shao, Steven P. Statistics for Business and Economics, Columbus OH: Charles E. Merrill Publishing Company, 1976.

25. Devore, Jay L. Probability and Statistics for Scientists and Engineers, Monterey CA: Brooks/Cole Publishing Company, 1987.

Vita

Captain Albert A. Larkin ████████████████████████████ ██████, and graduated from high school in Braintree, Massachusetts in 1980. After graduating from Norwich University in Vermont with a degree in Electrical Engineering in 1984, he entered the Air Force and was assigned to the Strike Systems Program Office at Wright-Patterson AFB where he was assigned to the Low Altitude Navigation and Targeting Infrared System for Night (LANTIRN) program. Later he was program manager for the Infrared Search and Track System program until entering the School of Systems and Logistics, Air Force Institute of Technology, in May, 1989.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE September 1990	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE MISLEADING GRAPHICS: CAN DECISION MAKERS BE AFFECTED BY THEIR USE?	5. FUNDING NUMBERS

6. AUTHOR(S)

Albert A. Larkin, Capt, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology WPAFB OH 45433-6583	8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GSM/LSY/90S-18

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This thesis investigated how Air Force decision makers can be affected by misleading graphics. A literature review revealed the existence of criteria for creating high-integrity graphs, as well as style guides for formatting and procedural techniques. By violating the criteria, misleading graphs may be produced which misrepresent the underlying data. Experiments were conducted on 63 Air Force Institute of Technology short-course students. Graphs typically used in the analysis of cost performance reports were constructed in a way to violate one of the criteria for the experiments. Using a t-test, it was demonstrated that Air Force decision makers can be misled by graphs that violate the criteria at a level of significance of less than .0001. Also, a sampling of graphs from throughout Air Force Systems Command revealed that program managers and cost analysts were creating graphs that violated the criteria.

14. SUBJECT TERMS Graphs, Cost Analysis, High-Integrity, Decision Making, Misleading, Criteria, Thesis, Sbw	15. NUMBER OF PAGES 131
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL